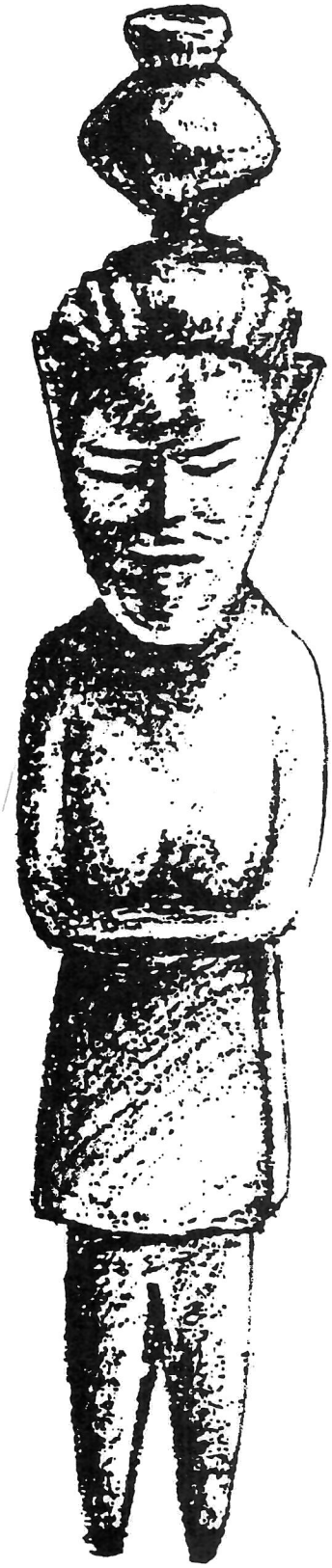
MOTHER TONGUE

JOURNAL OF THE ASSOCIATION FOR THE STUDY OF LANGUAGE IN PREHISTORY

Issue 1, December 1995



CONTENTS

1 Introduction

Articles

3 Basque and Dene-Caucasian: A Critique from the Basque Side: *R. L. Trask*

Discussants:

83 *Paul K. Benedict*
84 *John D. Bengtson*
104 *Václav Blažek*
111 *Eric de Grolier*
115 *José Ignacio Hualde*
120 *William H. Jacobsen, Jr.*
143 *Edwin G. Pulleyblank*
149 *Merritt Ruhlen*
157 *Vitalij Shevoroshkin*
159 *Étienne Tiffou*
163 *Roger W. Wescott*
165 *Xabier Zabaltza*

172 Response to Comments: *R. L. Trask*
202 Canaanite Vocabulary in Bengali and in Some Other IE Dialects of India *Liny Srinivasan* and *Cyrus Gordon*
207 The Concept of Proof in Genetic Linguistics: *Joseph H. Greenberg*

Book Reviews

217 *The Emergence of Homo Sapiens and His Languages in Tropical Asia* by Wilfried W. Schuhmacher, Juan R. Francisco, and F. Seto Reviewed by *Robert Blust*
219 *Indo-European and the Nostratic Hypothesis*, by Allan R. Bomhard Reviewed by *Igor M. Diakonoff*
223 *The Great Human Diasporas: The History of Diversity and Evolution* by Luigi Luca Cavalli-Sforza and Francesco Cavalli-Sforza Reviewed by *Daniel McCall*

Editorials

229 Editorial I: Surveying the Damage
233 Editorial II: Afrasian Gets, Better, Better, and Deeper

Debate and Discussion

235 More on the Austric Hypothesis and Austronesian's Inclusion *La Vaughn H. Hayes*

OFFICERS OF ASLIP

(Address appropriate correspondence to each.)

President:	Harold C. Fleming 16 Butman Avenue Gloucester, MA 01930 U.S.A.	Telephone: (508) 282-0603	Co-editor of *Mother Tongue*
Vice President:	Allan R. Bomhard 88-B Queen Street Charleston, SC 29401 U.S.A.	Telephone: (803) 720-8531 E-mail: bomhard@aol.com	Co-editor of *Mother Tongue*
Secretary:	Anne W. Beaman P.O. Box 583 Brookline, MA 02146 U.S.A.		

BOARD OF DIRECTORS

Ofer Bar-Yosef
Harvard University

Ron Christensen
Entropy Limited

Frederick Gamst
University of Massachusetts

John Hutchison
Boston University

Mark Kaiser
Illinois State University

Mary Ellen Lepionka
Cambridge, MA

Philip Lieberman
Brown University

Daniel McCall
Boston, MA

Roger Wescott
Southbury, CT

COUNCIL OF FELLOWS

Raimo Anttila
UCLA (USA)

Luca Luigi Cavalli-Sforza
Stanford University (USA)

Igor M. Diakonoff
St. Petersburg (Russia)

Aaron Dolgopolsky
University of Haifa (Israel)

Ben Ohiomamhe Elugbe
University of Ibadan (Nigeria)

Joseph H. Greenberg
Stanford University (USA)

Carleton T. Hodge
Indiana University (USA)

Dell Hymes
University of Virginia (USA)

Sydney Lamb
Rice University (USA)

Winfred P. Lehmann
University of Texas (USA)

Karl-Heinrich Menges
University of Vienna (Austria)

Colin Renfrew
Cambridge University (UK)

Vitaly Shevoroshkin
University of Michigan (USA)

Sergei Starostin
Academy of Sciences of Russia (Russia)

© 1995 by the Association for the Study of Language in Prehistory
ISSN: 1087-0326

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

INTRODUCTION TO *MOTHER TONGUE*: THE JOURNAL

Welcome aboard, everyone! Some nine years (almost to the day) after our first epistle was dispatched, we have transcended the newsletter stage and developed into a proper journal. For more "newsy", things we still generate small quarterly newsletters, but the most serious scientific discussions will take place herein.

Our general format is frankly copied — in admiration — from *Current Anthropology*. Recognizing that so-called peer review is a matter of paramount importance to some colleagues, we have instituted such procedure herein. But also noting that peer review is often a secretive and blatantly unfair process, not only in journals but also in funding agencies, we will follow *Current Anthropology*'s famous CA*Treatment in general and often in detail. Our procedure is called MT*Treatment.

We value truth seeking and hypothesis formation, but also critical examination and the attempt to falsify. These values manifest themselves as rebellion, on occasion, when we perceive established institutions and comfortable consensa as blocking the road to truth. We judge such obstacles to be present and uncommonly serious in current "Americanistics" (both historical linguistics and archeology). Some established thinking looks more like Maginot Line reaction than truth seeking, and one perceives a tacit Fortress (Indo-)Europa around the Indo-European language family, lest its ancestor be bundled with unsavory aliens in a genetic linguistic class.

We are fundamentally interdisciplinary, with our most salient divisions being biogenetics, paleoanthropology, archeology, and historical linguistics. The last tends to dominate discussions because the fiercest general resistance to our values and goals comes from linguistics. Our collection of interests has been most aptly named "the emerging synthesis" by Cambridge archeologist, Colin Renfrew. The dovetailing and intermingling of hypotheses from the divisions offer powerful insights into the deeper prehistory of humankind. Yet the conflicts between them often force re-thinking of the conclusions reached in one division. Presently, for example, biogenetic and linguistic hypotheses are at loggerheads about the peopling of eastern Asia, due to conflicts in taxonomic theories.

We try to write in ordinary, or at least unspecialized, English. While this is a source of criticism — our alleged "flakiness" troubles some proper colleagues —, it is necessary for communication between divisions. Both modern biogenetics and linguistics are perceived as impenetrable jargons by colleagues in adjoining fields, while archeology strains to become unintelligible too. The model here seems to be austere scientific publications where only initiates can understand the mysteries within specific disciplines. As a consequence of this linguistic specialization, scholars in adjoining fields are often remarkably ignorant of their peers in the next room, down the hall, so to speak. This whole matter of scientific specialization, while often lauded as a necessary evil, can in fact lead to fruitlessness. It depends on the problem being worked on. In the saying: "He knew more and more about less and less, until he knew everything about nothing", we behold some uncomfortable truth. Talking to each other in the hallway, as it were, may help each of us grasp the general problem and what progress we all are making.

The general problem is easy enough to state. We reckon that human language is closely related to the advent of more complex cultures and to anatomically modern *Homo sapiens*. Currently, the leading hypotheses with these assumptions as background propose that *Homo sapiens sapiens* and/or immediate predecessors "invented" human spoken language, intensified human social capacity, expanded human knowledge immeasurably, and (as a most impressive competitor for resources) spread around the Old World, eliminating or absorbing pre-modern humans in the process. One corollary of this is that all known human spoken languages are genetically related to each other as descendants of that first invention — Ur-Human or Proto-Language. One test of that is to show a taxonomy of human languages — convincingly to linguists — which makes possible a universal family tree and ultimately the reconstructions of major cultural events associated with the evolution of modern people. Another corollary is that the complex evolution of physical humans — population movements and shared mutations — can be figured out and related to a universal family tree which can be dated and located to its roots. Finally, the tests of these theories can be made through archeological discoveries — eventually.

An African homeland, most likely in eastern Africa, is favored for the roots of both trees. Even the probable dates of emergent language, culture, and physique — circa 100,000 years ago — are mentioned frequently. By a growing consensus, Southeast Asia is favored as 2nd or 3rd archaic locale, a staging area in the great diaspora before the almost equally old settlement of Australasia. Much of this scenario has archeological bases, primarily in

East Africa, the Levant, and Australia, but rather later in Europe.

Very strong opposition to this emerging synthesis comes from paleoanthropology and a minority of biogeneticists. Oddly enough, linguists are generally favorable to much of this synthesis, yet remain fervently opposed — supposedly in principle — to any demonstration of a universal family tree of languages. We can call this the linguist's Split Brain Syndrome; the left hand states that all human languages are probably related, as the right hand flatly rejects this.

The goal of our enterprise is to seek the truth as it pertains to the emerging synthesis about modern human origins. *Mother Tongue* is not committed to any single proposition, while we obviously view the whole scenario quite favorably. We do not know how far along we will get towards the acceptance, falsification, or modification of general theory pictured above. There will no doubt be surprises. New excavations always have the power to falsify the most appealing hypotheses. Someone may break through one of the Maginot Lines. Linguists may integrate their brains, softening their rigor and moving towards their left hands. Who knows?

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

BASQUE AND DENE-CAUCASIAN: A CRITIQUE FROM THE BASQUE SIDE

R. L. Trask
School of Cognitive and Computing Sciences
University of Sussex, Brighton BN1 9QH, England
larryt@cogs.susx.ac.uk

1. Introduction

For generations, the genetic isolate Basque has attracted the attention of linguists hopeful of finding a relative for it somewhere. By the eighteenth century, several Basque writers were arguing that Basque must be descended from the ancient Iberian language of Spain, an idea that was later picked up and popularized by Wilhelm von Humboldt. Only with the decipherment of the Iberian script in the twentieth century could the proposal be investigated, however, and then it did not take long for Antonio Tovar and Luis Michelena to show that Basque was of no assistance whatever in reading the Iberian texts, and hence that Iberian could not be an ancestral form of Basque. In contrast, the far more fragmentary remains of the ancient Aquitanian language of southwestern Gaul have proved to be transparently Basque in many respects, and few vasconists now doubt that Basque is the more-or-less direct descendant of Aquitanian (see especially Michelena 1954).

Meanwhile, there was no shortage of alternative proposals. The distinguished vasconist Hugo Schuchardt pursued connections between Basque and African languages, especially Berber, an idea which has been continued by others, most notably Hans Mukarovsky. Other linguists have attempted to demonstrate links between Basque and virtually every language or family spoken or recorded in the Old World: Pictish, Etruscan, Minoan, Sumerian, Indo-European, Uralic, Paleo-Siberian, Austroasiatic...the list is endless. None of this work has been convincing, and all of it is couched entirely in what I have elsewhere called the "Bongo-Bongo" approach: "I've got a few Basque words here that look quite a bit like some words in Bongo-Bongo." For surveys of all these efforts, see Trask (forthcoming *b,* forthcoming *c*).

Unquestionably, the most popular candidates for connections with Basque, however, have been the Caucasian languages, both North and South. Attempts at linking Basque genetically to some or all of the Caucasian languages have been underway for nearly a century, and the volume of work in this area probably exceeds all other work on Basque genetic connections put together. This is not because of any great success in these investigations, but only because of typological similarities. Basque shares its ergative morphology and its elaborate system of verbal agreement in varying measure with most of the Caucasian languages, and the common presence of these non-Indo-European characteristics has been enough to persuade any number of linguists that there must be a connection there to be discovered — a dangerous assumption, of course, since typological resemblances have rarely proved to be of much assistance in identifying genetic relations.

On the whole, the work on Basque and Caucasian has been considerably more sober and careful than most of the other lines of inquiry. It was begun in a small way by Schuchardt, who was chiefly interested in finding a North African connection for Basque, but he occasionally, as in Schuchardt (1913), cited some Caucasian parallels. The Dutch linguist C. C. Uhlenbeck pursued the Basque-Caucasian connection throughout his career, for example in Uhlenbeck (1923, 1924, 1940-41, 1946, 1947). The Italian linguist Alfredo Trombetti produced an entire book (1925) claiming a long list of Basque-Caucasian cognates. The Russian linguist Nikolai Marr, in the days when he was no longer respectable, published several articles comparing Basque and Caucasian. The French Caucasianist Georges Dumézil devoted a chapter of his 1933 book on North Caucasian languages to citing a number of supposed cognates with Basque. The French vasconist René Lafon produced a long series of papers arguing for a Basque-Caucasian genetic link and proposing some dozens of cognates (Lafon 1933, 1944 [appendix to vol. 1], 1948, 1951, 1952a, 1952b, 1957, 1967, 1968). The Norwegian Caucasianist Hans Vogt pursued the question in two papers and proposed a modest list of cognates (Vogt 1942, 1955), though his conclusions are more negative than otherwise. Finally, the German linguist Karl Bouda, the most enthusiastic of all the proponents of a Basque-Caucasian link,

after surveying the entire earlier literature and devising some further comparisons of his own, put the seal on the whole enterprise by presenting nearly 500 putative cognates in a series of papers (Bouda 1948, 1949, 1951, 1952).

Though most of these investigations are mercifully free of the kind of fantasizing that characterizes so much of the work I have alluded to above, the blunt fact is that they do not measure up to the standards normally expected in establishing genetic relationships. With the partial exception of Lafon (1948), which at least attempts (unconvincingly) to identify some systematic correspondences between Basque and Georgian sibilants, all of this work remains at the level of butterfly collecting: the investigators achieve nothing, and indeed attempt nothing, beyond compiling lists of Basque words and morphemes which bear some kind of resemblance to words and morphemes in one Caucasian language or another. But, with some thirty-eight highly divergent languages to play with, they could hardly fail to find such resemblances, particularly since the Caucasian "cognates" they cite are in nearly all cases merely items found in some particular language, items which cannot be shown to have existed in any version of Proto-Caucasian .

Indeed, there are reasons to suspect that the Caucasian languages (especially the northern ones) are *a priori* most implausible candidates for being relatives of Basque. For one thing, roots in North Caucasian languages are typically very short, often no more than a single consonant, while Basque roots are typically much longer and never consist of a single consonant. Furthermore, Basque has a very modest consonant system, and, moreover, the consonant system of pre-Basque was even more impoverished, consisting of no more than sixteen consonants and possibly of as few as eight (see Michelena 1957a, 1977; Trask 1985, forthcoming c). Some North Caucasian languages, in complete contrast, have the largest inventories of consonants on the planet. The recently extinct Ubykh had no fewer than eighty consonants, while many others exhibit between fifty and seventy. Nor is work in reconstruction obviously simplifying this picture: according to Catford (1991:265), a recent reconstruction of the Proto-Lezgian subgroup of Northeast Caucasian posits the startling total of 101 consonants, while the first attempt at reconstructing Proto-North-Caucasian has provisionally set up the astounding total of 180 consonants. Hence it would appear that *any* attempt to relate Basque to North Caucasian would have to assume either that the Caucasian languages have undergone a comparatively recent explosion of their consonant system or that Basque has undergone a catastrophic meltdown of its system on a scale not paralleled elsewhere. Even the dramatic reduction of the Proto-Indo-European obstruent system exhibited by Tocharian, often singled out as the most profound system collapse ever discovered, was not of this magnitude, and was moreover accompanied by the introduction of some new consonants.

Nevertheless, this body of work was received with enthusiasm in some quarters. The Spanish linguist Antonio Tovar apparently accepted the proposed Basque-Caucasian unity without hesitation in his various works on Basque (*e.g.*, Tovar 1950, 1959), and it is not difficult to find other linguists who take a similarly favorable view. But the proposal also encountered some formidable opposition in the person of the great Basque linguist Luis Michelena.

Now Michelena was in no way hostile to the idea of a Basque-Caucasian genetic link. Quite the contrary: by his own admission, he would dearly have loved to find some relatives for Basque, and he took a keen interest in all such work. He even contributed to a lexico-statistical study of Basque and Caucasian (Tovar *et al* 1961), and, in his (1950a) article, he went so far as to point out a Basque-Georgian parallel which had been overlooked. Nonetheless, he had a very clear understanding of what could be counted as evidence, and he did not find such evidence here.

In two reviews of Bouda's work (1950b, 1953), Michelena roasts the German linguist for playing fast and loose with the Basque data — in particular, for arbitrarily segmenting Basque words in order to extract the portions he wants to match, while airily dismissing the remaining material as ancient "prefixes" or ''suffixes''. In his (1964b) book, he complains that the Basque/Caucasian proposals rest on nothing more than random similarities, that they mostly display a shocking ignorance of Basque, and that they have succeeded in shedding no light at all on the prehistory of Basque. And, in an article published in 1968, by which time the work on Basque and Caucasian had largely dried up, Michelena reviews the whole body of such work. Though politely phrased, as always, this review is scathing, even devastating. In his considered opinion, no evidence of any significance at all can be extracted from this long list of publications. All those carefully compiled lists of putative cognates amount, in Michelena's view, to nothing more than lists of random similarities between Basque and one or another Caucasian language. Michelena is confident that an equally impressive list of "cognates" could be found between Basque and any sample of thirty-

odd Indo-European languages, if anyone were willing to go to the trouble of looking for them. He finds nothing resembling systematic correspondences: the rule seems to be that any Basque segment can be matched with any remotely similar Caucasian segment, at the whim of the author. Even the grammatical parallels offered mostly involve very short morphs of simple form, such as *-a*, *-n*, *-k*, or *-ra*, of the sort that one might expect to find in profusion in any highly inflected language, and these parallels never involve systematic alternations.

Michelena closes his review by suggesting that the linguists whose work he is surveying had started by simply *assuming* that Basque and Caucasian must be related, and that they had therefore proceeded merely to collect possible confirming instances, without attempting any sort of scrutiny of their work. Michelena's assessment of this work, I am confident, must be accepted by anyone who takes historical linguistics seriously.

Michelena's review effectively dismissed the entire Basque-Caucasian enterprise as something close to a total waste of effort, and for nearly two decades hardly anyone seems to have pursued the matter further, except that some of Lafon's Basque-Kartvelian work was republished in Russian as Lafon (1976). Then, in 1985, the Caucasian linguist V. A. Čirikba returned to the issue with a ten-page paper proposing a total of some 90 cognate words and grammatical morphemes between Basque and various North Caucasian languages (he excludes Kartvelian). It is difficult to tell whether Čirikba considers that he is presenting new work, or a summary of the earlier work, or a combination of both. On the one hand, he mentions the names of Bouda, Lafon, Trombetti, Dumézil, and Uhlenbeck, and he certainly repeats a large number of their proposed cognates. On the other hand, he cites no work by any of these scholars except for the irrelevant Uhlenbeck (1927) and the even more irrelevant Lafon (1976), he speaks in his brief introduction of the importance for his purposes of recent work in the reconstruction of North Caucasian proto-languages, and he presents a number of putative cognates which I have not seen elsewhere.

More recently still, several linguists have begun arguing for the existence of a vast and far-flung macrofamily which they call "Dene-Caucasian"; this construct sprawls across the planet from western Europe to the southwestern USA. The chief architects of this idea are John Bengtson, Merritt Ruhlen, and Vitaly Shevoroshkin, and these authors have maintained in a series of publications that Basque should be included in Dene-Caucasian — indeed, Bengtson and Ruhlen (1994:288) go so far as to assert that Basque may be "confidently" added to the family. Moreover, Bengtson goes further and argues that Basque, North Caucasian, and Burushaski (and no others) constitute a separate branch of Dene-Caucasian, a branch which he calls "Macro-Caucasian". These two related theses have been defended in a series of publications proposing some 300 putative cognates relating Basque to some or all of the other languages included in the grouping, all of which are presented below.

Indeed, not content with this ambitious undertaking, some of these linguists have gone further still. Bengtson and Ruhlen (1994) adduce Basque data in support of eight of the twenty-seven "global etymologies" which they have assembled — that is, they claim that these Basque items represent uninterrupted continuations in Basque of the lexical items present in the ancestral speech of all humankind ("Proto-World"), and to these Blažek (1992) adds one more. I include all these citations in my survey.

In this paper, I propose to scrutinize the plausibility of these "cognates" from the Basque side. I make no claims to expertise in any of the other languages adduced, and here I shall simply accept the authors' citations of these other languages without comment.

2. Preliminary remarks on the prehistory of Basque

In one of his articles, Bengtson (1994b:34) unhesitatingly accepts a few chance resemblances as evidence of loans into Basque from such implausible sources as Berber, Kartvelian, and even ancient Egyptian. In the same passage, however, he notes two Latin loans into Basque and then continues "[These] are among the few Latin words that have penetrated into Basque....But all this [*i.e.*, borrowing from all sources — RLT] has affected the basic vocabulary of Basque very little."

I find this statement astounding. Basque has been in the most intense contact with Latin and its Romance descendants for some 2000 years, and the impact of these neighboring languages upon the vocabulary of Basque has been profound — even overwhelming. The Latin and Romance words in Basque run into the thousands, and they penetrate almost every area of the lexicon: names of body parts (*kopeta* 'forehead', *matela* 'cheek', *hanka* and

zango, both 'leg', *kokots* 'nape' or 'chin', *kasko* 'skull', *ganga* 'palate', *tripa* 'stomach, belly', *diti* 'breast', *berga* 'penis', *potorro* 'vulva', *potro* 'testicle', *zola* 'sole [of the foot]'), names of plants (*pinu* 'pine', *gaztaina* 'chestnut', *lore* 'flower', *tipula* 'onion', *onddo* 'fungus', *arrosa* 'rose', *letxua* 'lettuce', *bago* 'beech', *biku* 'fig', *gerezi* 'cherry', *erika* 'heather', *arbola* 'tree'), names of animals *(ahate 'duck', antzara* 'goose', *katu* 'cat', *azeri* 'fox', *balea* 'whale', *izokin* 'salmon', *aingira* 'eel', *dordoka* 'turtle', *antxoa* 'anchovy', *arratoi* 'rat', *zamari* 'horse', *abere* 'domesticated animal'), names of tools and implements *(haizkora* 'ax', *para* 'spade', *mailu* 'hammer', *makila* '(walking) stick', *kaiku* 'wooden bowl', *baso* 'drinking-glass', *zamau* 'tablecloth', *soka* 'rope', *marmita* 'cooking-pot', *paru* 'pole', *orga* 'wagon'), nautical terms *(bela* 'sail', *gila* 'keel', *kai* 'wharf', *kabila* 'mast', *marinel* 'sailor', *batel* 'rowboat', *aingura* 'anchor', *erramu* 'oar', *portu* 'harbor'), articles of clothing *(zapata* 'shoe', *mauka* 'sleeve', *txapela* 'beret', *galtzak* 'trousers', *arropa* 'clothing', *alkandora* 'shirt', *gona* 'skirt'), common adjectives *(berde* 'green', *motz* 'short', *polit* 'pretty', *majo* 'nice', *merke* 'cheap', *garesti* 'expensive', *ziur* 'certain', *tonto* 'stupid', *pobre* 'poor', *franko* 'lots of'), names of natural phenomena *(zeru* 'sky', *mendebal* 'west wind', *aire* 'air', *laku* 'lake', *erripa* 'slope', *bortu* 'mountain pass', *irla* 'island', *muino* 'hill'), common verbs *(barkatu* 'forgive', *iskiriatu* 'write', *hautatu* 'choose', *saiatu* 'try', *aditu* 'hear, understand', *gustatu* and *laketu* 'please' (*i.e.*, 'like'), *deitu* 'call', *pentsatu* 'think', *geratu* 'stay', *tiratu* 'pull', *bota* 'throw'), and words from virtually every conceivable area *(musu* 'kiss', *kobre* 'copper', *diru* 'money', *bake* 'peace', *denbora* 'time', *gela* 'room', *zapatu* 'Saturday', *ordu* 'hour', *lama* 'flame', *putzu* 'well', *eliza* 'church', *ohore* 'honor', *moeta* 'kind, type', *magia* 'pod', *errege* 'king', *arrazoi* 'reason', *lukuru* 'usury, interest', *dorre* 'tower', *mila* 'thousand', *kale* 'street', *gauza* 'thing', *deus* 'anything', *merkatu* 'market', *kontra* 'against', *ondo* 'bottom', *kantu* 'song', *dantza* 'dance', *xaboi* 'soap', *funts* 'base, foundation', *leku* 'place', *lege* 'law'... I could go on like this for a long time. There are even a number of bound morphemes of Romance origin, notably noun-forming suffixes like *-ada, -aia, -antza, -ari, -dura, -eria, -eta, -men ~ -mendu, -tate* and *-zio.*

Naturally, the great majority of these loans have passed into Basque from the neighboring varieties of Romance: Castilian Spanish, Aragonese Spanish, and Occitan (especially Gascon). Very many of them, however, were clearly borrowed from Latin at a very early stage, before the occurrence of such major phonological developments in Latin as the palatalization of velars before front vowels and the reorganization of the five long and five short vowels into the seven-vowel system of western Romance. Here are a few examples showing the retention in Basque of the consonantism and vocalism of Latin: *lege* 'law' < LEGE, *errege* 'king' < REGE, *gerezi* 'cherry' < CERESEA, *bike* 'pitch' < PICE, *gela* 'room' < CELLA, *gisu* 'plaster' < GYPSU, *biku* 'fig' < FICU, (archaic) *lupu* 'wolf' < LUPU, *lukuru* 'usury' < LUCRU, *mundu* 'world' < MUNDU, *ingude* 'anvil' < INCUDE, *laku* 'lake < LACU, *bortu* 'mountain pass' < PORTU, *liburu* 'book' < LIBRU. Other words were clearly borrowed later and show the effects of the sound changes that transformed Latin into western Romance: *tipula* 'onion' < CEPULLA, *zeru* 'sky' < CAELU, *deitu* 'call' < DICTU, *putzu* 'well' < PUTEU. (Compare the modern Castilian Spanish forms of these words, in which <c> and <z> represent [θ]: *cebolla, cielo, dicho, pozo.*)

There is a further point, of central importance in searching for cognates of Basque words. The earlier scholars whose efforts were criticized by Michelena were working before the phonological prehistory of Basque had been elucidated; as a consequence, they neglected to take this history into account in their investigations, frequently with catastrophic consequences. In the 1950s and '60s, however, Michelena himself worked out the phonological history of Basque for the last 2000 years; he reconstructed the phonological system of pre-Basque in considerable detail for a date of about 2000 years ago, and he described the subsequent phonological changes which affected both the language as a whole and its individual dialects. This work is presented, with abundant documentation, in the magnificent volume which was first published in 1961 but which is here cited in the second, expanded edition of 1977. This book is one of the finest pieces of historical reconstruction that anyone could ever hope to read, and its conclusions are not seriously doubted by any vasconists. A summary of the history of Basque consonants is given in Michelena (1957a), and a survey of all the main points is provided in Trask (forthcoming c).

Curiously, the existence of this work does not appear to be widely known among historical linguists, and all too often one sees statements to the effect that "nothing is known about the history of Basque", or that "one can only work with the present-day forms of Basque words". Such statements could not be further from the truth. It is simply ludicrous to try to do any kind of historical work on Basque without taking Michelena's massive reconstruction into account — and yet that is exactly what has been done by the linguists whose work I shall be scrutinizing in this paper. All of them appear to be entirely unaware of the very existence of Michelena's work, and

they constantly operate with nothing but modern-day words and forms extracted from bilingual dictionaries and other secondary sources. As I shall demonstrate below, the consequences of this approach are just as catastrophic as they were for the earlier comparativists: the more recent workers persistently cite as data words which could not possibly have been present in Basque as recently as 2000 years ago, or words which, though present, could not have had the phonological forms which the comparativists require; above all, they constantly cite transparent loan words from Latin and Romance.

This is not the place to present the phonological history of Basque in detail; the interested reader is referred to the sources cited above. However, it will be necessary for me to present a few major points, in order that the reader may understand my reasons for rejecting large numbers of the proposed "cognates". When I refer to "native Basque words", I mean words which were present in the language 2000 years ago, at the time of the Roman invasion of the Basque Country.

[1] No native Basque word can begin with any of *p, t, k, d,* or *r.*

Note in particular that no native word begins with any voiceless plosive. Observe what happens in loans from Latin:

PAGO > *bago* 'beech'	PICE > *bike* 'pitch'
TEMPORA > *denbora* 'time'	TURRE > *dorre* 'tower'
CELLA > *gela* 'room'	CERESEA > *gerezi* 'cherry'
PARCERE > *barkatu* 'forgive'	CERTU > *gertu* 'certain'
PACE > *bake* 'peace'	*TASTARE > *dastatu* 'taste'

Latin initial voiceless plosives are consistently rendered by voiced plosives in Basque, simply because pre-Basque had no voicing contrasts, and the single series of word-initial plosives of pre-Basque appear in the Basque of historical times as voiced plosives. At some later stage, Basque acquired initial voiceless plosives, and some loans show these:

PORTU > *portu* 'harbor'	CATTU > *katu* 'cat'
CATENA > *katea* 'chain'	

In many cases, it is likely that this development was a consequence of the continuing influence of the neighboring Romance languages, all of which retain the initial voiceless plosives of Latin, and that the earlier loans were re-formed accordingly. Thus, while most varieties have *bago* 'beech' and *dorre* 'tower', some varieties have *pago* and *torre,* doubtless from Romance influence. A second factor is the sporadic but notable tendency for initial plosives to be devoiced if the following syllable contains a voiceless plosive. Thus, for example, in place of *bake* 'peace', some varieties have *pake,* and the forms *katea* 'chain' and *katu* 'cat' (just cited) are notably commoner than the more regular alternatives *gatea* and *gatu.* Such voicing assimilation is confined to loan words; native words like *bakar* 'lone, sole', *guti* 'not much, not many' and *betazal* 'eyelid' never show it, but then virtually the only native words of the relevant form are compounds: very few monomorphemic native words have plosives in the first two syllables, and those that do (like *gogo* 'soul', *gudu* 'combat', *gabe* 'without', *bide* 'road') usually have only voiced plosives. Basque words beginning with voiceless plosives, or even with *d-*, are thus *always* loan words or recent formations.

Latin and Romance words with initial *r-* always acquire a prothetic vowel:

ROSA > *arrosa* 'rose'	RIPA > *erripa* 'slope'
REGE > *errege* 'king'	ROMA > *Erroma* 'Rome'

[2] No native Basque word can begin with any consonant cluster at all.

Observe that initial clusters in Latin loans are always reduced in one way or another, or else a prothetic vowel is added:

FRONTE > *boronde* 'forehead'
CRUCE > *gurutze* 'cross'
PLUMA > *luma* 'feather'
GRANU > *garau* 'grain'
*CLETA > *gereta* 'rustic gate'
FLAMMA > *lama* 'flame'
SPATHA > *ezpata* 'sword'
PLACET > *laket* 'be pleasing'

[3] No native Basque word could have contained *m* in the pre-Basque period, because there was no */m/* in the language.

This is typologically unusual (though not unparalleled), but there appears to be no single case in which an */m/* can be confidently reconstructed for any native Basque word, and the frequent */m/* of modern Basque is found mostly in loan words and in "expressive" formations — onomatopoeic items, phonesthetic words, nursery words, and the like. In native words, Basque *m* derives from several sources: initial *m-* derives from **b-*, particularly under the influence of a following nasal (hence *mihi* 'tongue' < **bini;* note also cases like *mika* 'magpie' < **bika* < PICA) and *magi(ñ)a* 'pod' < **bagina* < VAGINA); medial *-m-* derives either from **-nb-* (hence *seme* 'son' < **senbe,* attested in Aquitanian as *SEMBE),* from **-b-* under the influence of a following nasal (hence regional *imiñi* 'put' < *ibeni;* note also cases like *zamau* 'tablecloth' < SABANU), or rarely from **-n-* preceded by *u* (hence *zumar* 'elm' < **zunar;* note also cases like *kuma* 'cradle' < CUNA).

[4] Except in the eastern dialects Z and R, no native Basque word can contain a cluster of *n* or *l* followed by a voiceless plosive.

Except in the eastern dialects, plosives were uniformly voiced in this position. Hence we find common *alde* 'side' but eastern *alte*; common *lagundu* 'help' but R *laguntu*, Z *lagüntü*. Latin loans show this process very well: TEMPORA > *denbora* 'time'; ALTARE > *aldare* 'altar'; INCUDE > *ingude* 'anvil'; SANCTU > *saindu* 'saint'.

[5] Pre-Basque intervocalic *-n-* was categorically lost.

There are only two minor and sporadic exceptions: original **-n-* is sometimes retained as palatal *-ñ-* after *i*, and occasionally retained as *-m* after *u*. Note the consequences of this in Latin loans; for the unexpected *h* in some words, see below:

BALLAENA > *balea* 'whale'
CATENA > *katea* 'chain'
ANATE > *ahate* 'duck'
ORGANA > *orga* 'wagon'
MANICA > *mauka* 'sleeve'
SENAPE > *ziape* 'mustard'
VAGINA > *magia* ~ *magiña* 'pod'
CORONA > *koroa* 'crown'
MONETA > *moeta* 'kind, sort'
HONORE > *ohore* 'honor'
SABANU > *zamau* 'tablecloth'
GRANU > *garau* 'grain'
L1NU > *liho* ~ *liño* 'flax'
CUNA > *kuma* 'cradle'

Compare the treatment of Latin geminate *n*:

ANNONA > *anoa* 'provisions'
IOHANNES > *Joanes* 'John'

The loss of *-n-* left behind nasalization of the adjacent vowels. This nasalization usually survives today in the eastern dialects Z and R. Elsewhere, nasalization was variously either lost or reinterpreted as a following *n*:

**zani* 'watchful' > Z R *zãi*, cornmon *zai* ~ *zain*
**zunai* 'hay, fodder' > *zuhai* ~ *zuhain* ~ *zumai*

**anatzi* 'forget' > Z *ãhãtze*, R *ãtze*, common *ahantzi ~ ahatzi*
**arrani* 'fish' > *arrai ~ arrain*

In word-formation, many of these words retain their *n* when occurring as the first element, because of other changes which prevented the *n* from being intervocalic:

**ardano* 'wine' > *ardao ~ ardo ~ ardũ*; combining form *ardan-*, as in *ardandu* 'ferment', *ardantza* 'vineyard'
CATENA 'chain' > *katea;* CF *katen-*, as in *katenbegi* 'link of a chain'
**arrani* 'fish' > *arrai;* CF *arran-*, as in *arrantzale* 'fisherman'
ORGANA > *orga ~ orgã* 'wagon'; CF *organ-*, as in *organbide* 'cartpath'

All this makes it generally easy to recover lost intervocalic nasals.

[6] Pre-Basque intervocalic *-l-* was categorically changed to *-r-*.

Again, observe the effect upon Latin loans:

GULA > *gura* 'desire'
ASCIOLA > *haizkora* 'ax'
ANGELU > *aingeru* 'angel'
PADULE > *madura* 'water meadow'
COLU > *goru* 'distaff'
SOLU > *zoru* 'ground', *soro* 'field'
PALU > *paru* 'pole, stake'
MILUU > *miru* 'kite' (bird)

But:

ANGELLU > *angelu* 'soil'
CELLA > *gela* 'room'
SELLA > *zela* 'saddle'
BALLAENA > *balea* 'whale'

Again, the original *l* is usually preserved in word-formation:

**gali* 'wheat' > *gari;* combining form *gal-*, as in *galburu* 'head of wheat', *galgorri* 'a variety of wheat'
ASCIOLA > *haizkora* 'ax'; CF *haizkol-*, as in *haizkolbegi* 'the hole in the axhead for the shaft'

[7] The pre-Basque aspiration was not etymological, but was merely a suprasegmental feature.

This is a fundamental point, and one which I cannot stress too strongly: with only a tiny handful of possible exceptions, the Basque aspiration does not continue an earlier segment. In Michelena's interpretation, the aspiration was associated in pre-Basque with the position of the word-accent. The aspiration survives today in most of the French Basque varieties; south of the Pyrenees, it has been lost. In the standard orthography, <*h*> is written wherever the French Basques have it word-initially or between vowels, but not elsewhere. The non-etymological nature of the aspiration can readily be seen in Latin loans; here I cite the forms from the dialects which retain the aspiration:

ARENA > *harea* 'sand'
HONORE > *ohore* 'honor'
OPTARE > *hautatu* 'choose'
LEONE > *lehoin* 'lion'
ANNONA > Z *anhua* 'provisions'
BACILLA > *makhila* 'stick'
ARMA > *harma* 'weapon'
ANATE > *ahate* 'duck'
ASCIOLA > *haizkora* 'ax'
SOLU > *sorho* 'field'
PIPER > Z *phiper* 'pepper'
CERTU > *gerthu* 'certain'

Note that cases like *ohore* derive from earlier forms like **onhore,* by the loss of intervocalic *n,* which is not prevented by the aspiration, confirming that *h* was not a consonant in pre-Basque. Even in the aspirating dialects, there is considerable variation as to the presence and even the location of the aspiration: for example, we find both *sartu* and *sarthu* 'enter'; *etorri* and *ethorri* 'come'; *osin* and *hosin* 'deep place in a river'; *hilargi* and *ilhargi* 'moon'; *hilerri* and *ilherri* 'cemetery'. A particularly good example of this *is golko* 'bay, gulf', a loan from Latin COLPU, for which all four of the variants *golko ~ golkho ~ kolko ~ kholko* are well attested in the aspirating dialects. Some of the authors whose work I shall be examining below erroneously treat the Basque aspiration as continuing an earlier segment and try to find cognate segments for it in other languages, an utterly futile procedure.

Reflecting its origin as a suprasegmental, the aspiration is subject to severe constraints:

(1) It can only occur on a syllable which otherwise begins with a vowel, a voiceless plosive, a liquid, or *n.*
(2) It cannot occur later than the onset of the second syllable.
(3) There can be no more than one aspiration per word.

With these phonological preliminaries out of the way, we can now turn to an examination of the evidence proposed for relating Basque to a putative "Dene-Caucasian" family. The reader is particularly invited to note the very great frequency of arbitrary segmentations provided by the authors, a practice which I shall be discussing at the end of the paper.

3. The comparisons

The Basque words and morphemes which have been adduced in comparisons are listed below in alphabetical order according to their forms in the standard orthography; the nonexistent forms adduced are also listed in alphabetical order. Within each entry, all citations, transcriptions, segmentations, glosses, and annotations are those of the original authors; these I have reproduced as faithfully as possible, with certain exceptions: where a proto-Caucasian form is adduced, I have not troubled to repeat the forms listed from individual languages; I have ignored the data from Sino-Tibetan and Na-Dene; and I have included "Proto-Dene-Caucasian" reconstructions only where these are directly relevant to my discussion. Where there are minor variations in transcription among different publications, I have preferred Bengtson's versions, and, among Bengtson's papers, Bengtson (199lc). Abbreviated references are as follows: B (John Bengtson), BB (Václav Blažek and John Bengtson), BK (Václav Blažek), BR (John Bengtson and Merritt Ruhlen), C (Vjačeslav A. Čirikba), R (Merritt Ruhlen), SR (Sergei Starostin and Merritt Ruhlen). 'PS' indicates a postscript to a paper. Numbers in square brackets are the numbers given in the original publications, where these exist. For example, (B94b [33]) denotes item 33 from Bengtson (1994b).

Note that Čirikba (1985) compares Basque only with North Caucasian, and says nothing about Dene-Caucasian or any other wider grouping. I do not know what his views might be on this issue. However, his work is frequently cited with approval by Bengtson; Bengtson (1993:3) calls Čirikba's paper "a giant step forward"; and Bengtson (199la: 81) explicitly accepts Čirikba's work as essentially valid. I trust, therefore, that I am doing no one a disservice by including Čirikba's work here.

The following additional abbreviations are used:

Languages:

Abkh:	Abkhaz
Bq:	Basque
Bur:	Burushaski
Cauc:	Caucasian
ECauc:	Proto-East-Caucasian
PAA:	Proto-Abkhaz-Adyghe

PAT:	Proto-Abkhaz-Tapantan
PDC:	Proto-Dene-Caucasian
PNC:	Proto-North-Caucasian
PND:	Proto-Nakh-Daghestan
PNEC:	Proto-Northeast-Caucasian
PW:	"Proto-World"
PY:	Proto-Yeniseian
W:	Werchikwar dialect of Burushaski
WCauc:	Proto-West-Caucasian

Dialects of Basque:

Aezk	Aezkoan
B	Bizkaian
G	Gipuzkoan
HN	High Navarrese
L	Lapurdian
LN	Low Navarrese
R	Roncalese
Sal	Salacenco
Z	Zuberoan

Among the sources of information which I cite most frequently, Azkue (1905) and Lhande (1926) are the two great scholarly dictionaries of Basque which we possess, Agud and Tovar (1988-) is a comprehensive etymological dictionary of Basque, and Corominas and Pascual (1980) is the standard etymological dictionary of Spanish, which includes careful treatment of words shared by Basque and Iberian Romance.

In the Basque orthography, <z>, <s>, and <x> represent three contrasting voiceless sibilants (laminal, apical, and palato-alveolar, respectively) and <tz>, <ts>, and <tx> are the corresponding affricates; <tt>, ll>, and <ñ> are a voiceless palatal plosive, a palatal lateral, and a palatal nasal; <h> represents [h] in most French Basque varieties but zero elsewhere.

[1.] Bq **-a** (article)

Bq **a** (demonstrative pronoun and definite article): PAA ***a** (demonstrative pronoun) (C85 [18])

The Basque article *-a* is, in origin, merely the absolutive case form of the distal demonstrative stem *har-;* today *har-* (as *a)* survives in the absolutive singular as a demonstrative only in the Bizkaian dialect, all other dialects having replaced it there with the suppletive form *hura* (though the stem *har-, hai-* remains in all other case forms, including the absolutive plural). In all dialects, however, *-a* (phonologically bound to a preceding element) has developed into a definite article, and there is now only a suprasegmental distinction between, for example, datives *gizon hari* 'to that man' and *gizonari* 'to the man'. In Bizkaian, the same is true of absolutives *gizon a* 'that man' and *gizona* 'the man' (though some varieties of Bizkaian have introduced a pleonastic form *a gizon a* 'that man'). Moreover, this development cannot even be particularly ancient in Basque: there is no trace of it in Aquitanian, and the glossary of Basque words elicited from native speakers by the French pilgrim Aimery Picaud in the twelfth century cites about half the nouns without the article, suggesting that the article was then only beginning to come into use. The article is therefore a development entirely within Basque of the post-Roman period and cannot serve as a basis of remote comparisons.

[2.] Bq **aberats** 'rich'

Bq a-**berats** 'rich' : PNC ***berčV** 'rich' : Bur **bārIš** 'things' (B9lb [56]), (B9lc [9], [42])

Basque *aberats* is almost certainly a derivative of *abere* '(domesticated) animal', reflecting the identification of wealth with livestock in a traditionally pastoral society, and *abere* itself is derived by Michelena (1977:226) from Latin HABERE 'have'.

[3.] Bq **abets** 'voice'

Bq **abe(t)s** 'voice': Abkh **a-bž'ə** 'voice' (C85 [55])

This entry is simply beyond belief. The Basque word for 'voice' is *ahots,* a derivative of *aho* 'mouth'. What C is apparently citing in its place is one of the fantastic neologisms coined by the Basque nationalist writer Sabino de Arana in the late nineteenth century. Arana made strenuous attempts to construct and introduce "genuine" Basque words in place of the thousands of Latin and Romance loan words in Basque, but his knowledge of Basque word formation was imperfect and bedeviled by absurd speculations; his creations were rarely well-formed, and only a handful have ever found a place in the language. The one cited here by C was assembled by Arana from *aho* 'mouth' and *eresi* 'dirge, song, elegy, poem' with his usual degree of attention to phonological detail.

[4.] Bq **adar** 'horn', 'branch'

Bq **adar** 'branch (of a tree)', soin-**ada** 'bodily extremity' (from *soin* 'body') : Abkh **ada** 'root', possibly also **ada** 'nerve', **ada**-š 'vein'; PND ***tɔ̄mɦi** 'vein' (C85 [77])

Bq a-**dar** 'horn; branch' : PNC ***tˡVrV** 'horn': Bur **-ltūr, tur** 'horn' (B9lb [3]) (B9lc PS1) (B9ld) (B93) (B94b [29])

Basque *adar* means both 'branch' and 'horn', and *soin-adar* (not **soin-ada*) is a transparent compound of it. But *adar* never has any meaning such as 'root' or 'nerve' or 'vein' (body part), though it is used for 'vein (of ore)' in mining. It seems to me that C is stretching the semantics too far here. Moreover, B differs from C in drawing a completely different set of parallels.

Michelena (1964b:139) notes the striking resemblance of the Basque word to Old Irish *adarc* 'horn', which has no convincing Indo-European etymology, and is inclined to impute both the Basque and the Irish words to some pre-Indo-European substratum.

[5.] Bq **adin** 'age'

Bq a-**din** 'age' : Bur **den** 'year' (B9lc [26])

Basque *adin* is the universal word for 'age' today, and it also means 'contemporary' in places. A potential difficulty, however, is that our earliest Basque texts use the word not only for 'age', but also for 'understanding, judgement', for 'conduct', and for 'season, time of life', 'mood', 'maturity', and 'culmination, state of perfection'. It is not easy to know how to interpret these facts, and some have suggested that the senses 'understanding, judgement' point to a derivative of **aditu,** discussed below. Though this leaves the final *-n* unexplained, a semantic development along the lines of 'understanding' > 'mature reflection' > 'maturity' > 'age' does not seem *a priori* implausible, though I would not endorse this idea myself. See Agud and Tovar (1988-) for an impressive list of similar-looking words in Afroasiatic and other languages, including an earlier citation of the Burushaski word adduced by Bengtson.

[6.] Bq **aditu** 'hear, understand'

Bq **adi** 'mind', 'idea', **adi**-tu 'understand' : Abkh a-**dər**-ra 'understand' (C85 [56])

The Basque verb *aditu* means 'hear', 'listen', 'understand', and it is a transparent borrowing from Latin AUDIRE 'hear', or more precisely from that verb's participle AUDITU (Latin verbs are always borrowed into Basque in their participial forms). The stem *adi* functions as an imperative meaning 'listen' or 'pay attention'. From this, a very sparsely attested noun *adi* has been derived which means variously 'attention', 'intention' or 'intelligence', but never 'mind' or 'idea' (Agud and Tovar 1988-).

[7.] Bq **agor** 'dry'

Bq **agor** 'dry', **igar** 'dry' : PAA ***R^{w}a-**, PAT ***ʕ^{w}a** 'dry' (C85 [41])

Bq a-**gor(r)** ~ i-**gar(r)** ~ i-**har(r)** 'dry': PNC ***ʔĭGG^{w}Vr-** 'dry, dry up' : Bur **qAqər** 'very old and wrinkled' : PY ***qVr$_{1}$-** ~ **qVl-** 'dry' (B9la [115])

Both C and B assume that these Basque words are related, but they are not. Basque *igar* is merely the western form of the word which appears in eastern dialects as *eihar,* and it shows the common western insertion of *g* to separate vowels in hiatus; the original form of the word must have been either **eiar* or **eiCar,* with a consonant now lost. The same pattern occurs in another word for 'dry' not cited: eastern *leihor,* western *legor.* It is just about conceivable that these two words are related in some way, but there is no parallel elsewhere in the language. The word *agor,* in contrast, occurs in precisely this form throughout the Basque Country, and hence is almost certainly not related to the other words.

[8.] Bq **agure** 'old man'

Bq a-**gure** 'old man', a-**gure**-tu 'grow old' : Bur **gūro** 'grayish' (B9la [127]), (B9lc [90])

It is generally thought by vasconists that the Basque word *agure* is a loan from Romance, though the precise source is uncertain. Most likely is the late Latin AVULU 'grandpa', vocative AVULE, since addressing an old man as 'grandpa' is almost everywhere a common practice. The problem is that Latin AVULE should have given **abure,* but fluctuation between intervocalic *b* and *g* before *u* is common in Basque (see **sagu** 'mouse' and **suge** 'snake' below). The transparent derivative *aguretu* should not be cited separately.

[9.] Bq **ahal** 'ability'

Bq **al** 'be able', 'ability' : Abkh **al**, **al**-ša-ra 'ability' (C85 [57])

The Basque form *al* is merely the western reduction of the more conservative *ahal,* which in isolation means 'ability, power'; only the compound verb *ahal izan* means 'be able'. The form of *ahal* strongly suggests a lost intervocalic consonant, and this consonant can in fact be recovered. The eastern dialects of Basque generally preserve, in the form of nasalized vowels, a record of an ancient intervocalic *n*. Though *ahal* itself shows no nasalization in these dialects, its compound *ahalge* 'shame' certainly does: Z *ãhãlke,* R *ãlke* (Michelena 1977: 461). This shows that the earliest recoverable form of the word must have been **anal,* a form not very similar to the proposed Caucasian cognates.

Interestingly, exactly the same matchup is cited by Trombetti (1925: Part 2, § 23), but is *rejected* by Bouda (1949:50) on the ground that the Abkhaz verb *alšara* has been wrongly analyzed: according to Bouda, the *l* here is merely a familiar preverb, and the root is *-ša-* alone.

[10.] Bq **ahizpa** 'sister (of a woman)'

Bq a-**hiz**-pa, a-**iz**-ta 'sister' (woman speaking): PNC ***ʔɨćć^{w}V** 'brother, sister': Bur **-Aǯo**, **-Aǯu** 'sister' (woman speaking), 'brother' (man speaking) : PY *b-**is(a)** or ***bis** 'brother, sister' (B9lb [43]) (B9lc PS1) (B9ld) (B93) (SR94 [35])

B *aizta* is a local variant of the otherwise universal *ahizpa.* The presence of a nasal vowel in eastern *a(h)ĩzpa* allows Michelena (1977:306) to reconstruct **anizpa,* destroying the proposed matches. It is possible that this word contains an element **an-* which it shares with *ahaide* 'relative' (< **anaide* ?) and *anai* 'brother' (of a man)' (< **annane* ?).

[11.] Bq **aho** 'mouth'

Bq a-**ho** 'mouth' (< Proto-Basque *a-x^{w}o) : PNC ***k'^{w}eɬʔV** 'mouth' : PY ***Xowe** 'mouth' : Sum **gu$_{7}$, kú** 'eat' (B9lc PS1, PS2 [22]) (B94b [36]) (SR94 [173])

As always, Basque *h* is not etymological, and Bengtson's **ax^{w}o* is no more than a flight of fancy. See the essay in Agud and Tovar (1988-).

[12.] Bq **ahuñe** 'kid'

Bq a-**huñe** 'kid' : PNC ***hI^{w}ɔ̃nxV** 'ram' (B9lc [119]) (B93)

For the word *ahuñe*, there is clear evidence of a lost nasal, and Michelena reconstructs a stem in **anu-* (1949a: 485; l950a: 454; 1977: 115), thereby destroying the match.

[13.] Bq **ahur** 'palm (of the hand)'

Bq a-**hur** 'fist, palm, hollow of the hand' : PNC ***k^{w}ĩlʕV** 'hand' (B9lb [5]) (B9lc [83]) (B93) (B94b [27])

The meaning 'fist' is certainly attested for the word *ahur ~ agur,* but all lexicographers and linguists agree that the primary meaning is 'hollow of the hand', 'palm'. Michelena (1971b:249-250) notes that *ahur is* found precisely in those eastern dialects in which the otherwise universal word *aurre* 'front part' is absent, and he therefore proposes that *ahur* is in origin merely a specialization of *aurre,* a proposal which he supports with comparable usages elsewhere in Basque and other languages. The final *e* of *aurre* has been acquired from its constant use in postpositions like *aurrean* 'in front of', in which the *e* is required for phonological reasons, and both words are derived from **a(h)ur* 'front'. Something very similar has occurred with the word which in the east is *(h)atz* 'trace, vestige' but which in the west is *atze* 'back part'. The match is therefore destroyed.

[14.] Bq **akain** 'tick'

Bq a-**kain** 'tick' : PNC ***q'(q')inʔV** 'louse' : Bur **khiin** 'flea' (B9la [99]) (B9lc [104]) (B9ld) (B93)

Basque *akain* also occurs as *lakain,* and the word is a loan from Romance, probably from Gascon *lagagno* (Michelena 1977:323). The initial lateral was no doubt lost because it was mistaken for the Romance article. See Corominas and Pascual (1980) under *legaña* for a discussion of the source of the word, and see the discussion under **itain** 'tick' below.

[15.] Bq **alaba** 'daughter'

Bq **alha**-ba 'daughter' : Hurrian **ela** 'sister' (B9lb [42])

Since no resemblance can apparently be found with Caucasian, Burushaski, or Yeniseian, I cannot see that a very vague resemblance in Hurrian adds anything to the argument. Agud and Tovar (1988-) catalogue a list of similar-

looking words which have been adduced in languages ranging from Guanche (the indigenous language of the Canary Islands) to the indigenous languages of Taiwan.

[16.] Bq **alderatu** 'separate, remove'

Bq **aldera**-tu 'separate, remove' : PNC ***ʔȭlt'ʷVr-** 'tear down, tear off' : Bur **īltUr**-as 'pull down, dismantle, pull to pieces, undo' (B91b [63])

Basque *alderatu* 'separate, remove' is assumed to contain an ancient stem **aldera-* of the same meaning; in fact, it is merely derived from *aldera* 'to the side', the regular allative case form of the noun *alde* 'side'.

[17.] Bq **alor** 'field'

Bq **alhor** 'field' : Sum **a.gàr, a.da.ar** 'field' (B9lc PS2 [32])

Since in much of the country Basque *alor* means specifically 'field ready for sowing', very many vasconists have seen this word as a derivative of *ale* '(individual) grain, seedcorn', which is semantically very plausible but which leaves the second element unexplained. The very vague Sumerian resemblance cannot be taken seriously.

[18.] Bq **ama** 'mother'

Bq **ama** 'mother' : Bur **māma** 'mother' : PY ***ʔama** 'mother' (SR94 [169])

The use of something like *mama* for 'mother' or 'breast' is so widespread, and so readily explicable in terms of babbling, that it cannot be counted as evidence for a genetic relationship.

[19.] Bq **amets** 'dream'

Bq a-**mets** ~ a-**mes** 'dream' : PNC ***Hnićʷ**V ~ ***Hćʷin**V 'night, dream' : PY ***sʌm-** 'sleep' (B9lc [54]) (SR94 [238])

As always, the *m* in *amets* cannot be original, and the word must derive from something like **anbets* or **abents,* either of which would destroy the match, such as it is. The second option is strongly supported by the widespread eastern variant *aments.*

[20.] Bq **ametz** 'gall oak'

Bq a-**mentz** 'a kind of oak tree' : PNC ***měnʒ**V 'oak' : Bur **meeš** 'bush ,shrub' (B93)

If this is a Basque native word, the *m* cannot be original; very probably we should reconstruct **anbetz,* or perhaps **abentz,* the second being supported by the rare attestation of regional forms like *amentx.* Agud and Tovar (1988-) catalogue several attempts at deriving the Basque word from Romance; these are very interesting, but not quite persuasive, and the word is attested in place names in the Basque Country but not elsewhere.

[21.] Bq **-antz** 'toward'

Bq **-antz** (postposition of arrival) : Abkh **-n-ʒa**, Abaza **-ʒa**, Ubykh **-ōnʒa** (postposition of arrival) (C85 [9])

The Basque suffix *-antz ~ -ontz ~ -untz* means 'toward', and there is good reason to suppose that *-untz* is the most conservative form. Precisely the same match is cited by Lafon (1951:§6) and dismissed by Michelena (1968:469) as insignificant. Internal evidence in Basque strongly suggests that this affix is a late accretion to the inflectional

morphology: it always follows the allative case-suffix, and the allative suffix itself appears to be partly a late development (see under **-ra** (allative) below).

[22.] Bq **apo** 'toad'

Bq **apho ~ aphu** 'toad' : Svan **apXw** 'frog' (B94b)
This is identified by Bengtson as a loan from Kartvelian into Basque. But the Basque word is variously *zapo ~ sapo ~ apo,* of which the first form has the earliest attestations, and it can hardly be separated from Romance forms like Castilian *sapo,* Aragonese *zapo,* Gascon *sàpou,* all 'toad'. It is probable that the unexpected *apo* derives from a mishearing of a Romance form like *los sapos* as *los apos* (Michelena 1977:292). Even if this is a native Basque word (which is possible), the primacy of *zapo* destroys the rather forlorn comparison with a single Kartvelian language.

[23.] Bq ****-ar**

Bq ***-ar** (plural suffix) : Abkh **-ar**, PND ***-rV** (C85 [4])

I am mystified by the assertion that Basque has or once had a "plural suffix" of the form *-ar*. No such suffix exists or can be reconstructed. In Basque noun phrases, the plural suffix is absolutive *-k*, oblique *-e-*, both probably deriving from **-g(e-)*; a few determiners and pronouns show a plural suffix *-(t)zu*. In finite verb forms, the absolutive plural marker is normally *-z*, extended to *-tza* or *-tzi* in a few localized forms, while the ergative plural marker is variously *-te, -de*, or *-e,* the first two certainly, and the third probably, derived from earlier **-de*. The common verb *ukan* 'have' uniquely exhibits an absolutive plural prefix *-it-* throughout its conjugation. Finally, the extremely irregular verb *izan* 'be' shows the following forms in the present indicative:

naiz 'I am'	*gara* 'we are'
haiz 'thou art'	*zara 'you* are'
da 'he/she is'	*dira* 'they are'

I can only assume that C has extracted the morph *-ar-* from two of these forms and unilaterally declared it a "plural marker". There can be no justication for such an arbitrary exercise, especially since this *-ar-* does not recur elsewhere, all forms of this verb other than the present indicative showing different irregularities.

[24.] Bq **ar** 'male'

Bq **ar(r)** 'male' : PNC ***Hir-k'^{w}V** 'man' : Bur **hiir** 'man, male' (B9la [131])

The unusual brevity of the Basque word has led to the citing of so many proposed "cognates" that even Agud and Tover (1988-) don't attempt to list them all. It appears that about half the languages on the planet have some kind of word for 'man' or 'male' containing a liquid.

[25.] Bq **ardo** 'wine'

Bq **ardo, ardao** 'wine' : Albanian **(h)ardhí** 'wine' : Armenian **ort^{h}** 'wine' (B9ld PS2)

This is identified by Bengtson as a loan from DC into the IE languages, but the Basque word, whose combining form is *ardan-* and which has a nasal vowel in the east (e.g., R *ardão*), is easily reconstructible as **ardano* (Michelena 1977: 151).

[26.] Bq **argi** 'light' (n.), 'bright'

Bq **argi** 'light', **argi**-zagi, hil-**argi** 'moon' : WC ***rəẏa** 'sun' = EC *wi-**ra̲GV** 'sun' (B94a [20]) (B94b [23])

The Basque word just means 'light, bright'; its attachment to the various regional words for 'moon' is clearly a late accretion (see ***hile-** 'moon' below). In any case, linguists have long suspected that this might be a loan word from IE, since a PIE root **arg-* 'shine, bright' is well attested (e.g., in Latin ARGENTU 'silver'), but no direct source for the loan (if it is one) can be identified.

[27.] Bq **arrain** 'fish'

Bq **arrain** 'fish' : Coptic **ra(a)me** ~ **rami** 'fish' (B94b)

This is somewhat surprisingly identified by Bengtson as a loan from ancient Egyptian into Basque. The Coptic resemblance was pointed out a century ago by Schuchardt, who reports, however, that the Coptic word denotes only a particular species of flatfish, and not 'fish' in general; on this, see Agud and Tovar (1988-).

[28.] Bq **arrats** 'evening'

Bq a-**rrats** ~ a-**rratx** 'night, evening' : PNC *H^{w}i-**ri̲nƛī** 'night, evening, yestereve' (B9la [71]) (B9lc [32]) (B9ld)

The Basque word means only 'evening'. Connections have been proposed for this word in an astounding number of languages; see Agud and Tovar (1988-) for a catalogue. The possibility cannot be ruled out that the word is merely a derivative of *arre* 'gray'.

[29.] Bq **arraultza** 'egg'

Bq a-**r(r)aultze** 'egg' : Bur i-**rič** 'kidney' (B9la [51]) (B9lc [31])

The widespread existence of variants with *n* in place of *l*, such as *arrauntza*, has induced many linguists to derive the Basque word from the verb *erron* ~ *errun* 'lay (eggs)'; something like **erra-(k)untza,* while not obviously right, is by no means implausible *(-kuntza is* a common noun-forming suffix).

[30.] Bq **arroda** 'wheel'

Bq a-**dor** 'wheel' : PNC ***t'wirV** 'wheel' (B93)

There is no such word as **ador* 'wheel'. What B is trying to cite is *adorra,* a severely localized metathesis of the more usual *arroda* 'wheel', a transparent loan from Romance (Occitan *roda,* Castilian *rueda,* from Latin ROTA). This a striking case of the replacement of an earlier Latin loan *(errota,* directly from ROTA) by a Romance loan.

[31.] Bq **aska** 'crib'

Bq **aska** 'trough, manger' : PNC ***ć'ä̲q'wV** 'scoop, vessel, spoon' : Bur **čuq** 'sieve, measure of grain' (B93)

See Agud and Tovar (1988-) for a survey of the numerous attempts at finding an etymology for the Basque word; they are inclined to reject all of these and to see the word as native, but note that the existence of an eastern variant *arska* muddies the waters considerably.

[32.] Bq **astigar** 'maple'

Bq a-**stigar** 'maple tree' : Batsbi (Cauc) **stager** 'maple tree' (B93)

The Basque word is variously *gaztigar* ~ *aztigar* ~ *astigar*, leading Michelena (1977:253) to interpret it as a compound of the element **cast-*, widely found in Romance tree names, and *ihar* 'maple'. In any case, the resemblance to Batsbi looks fortuitous.

[33.] Bq **asto** 'donkey'

Bq **asto** 'donkey' : PAA ***čʷVdə** 'donkey' (C85 [70])

A number of linguists have been unable to resist the temptation of deriving Basque *asto* from Latin ASINU of the same meaning (cf. Spanish *asno*). This, however, would be phonologically unusual, and anyway it overlooks the difficulty that the Zuberoan dialect, which tends to preserve consonant clusters better than other dialects, has *arsto*. The Basque lexicographer Azkue once suggested that *asto* might be derived from **hartz-to*, where *hartz* is 'bear' and *-to* is the ancient diminutive suffix; this proposal is semantically unexpected but phonologically impeccable, and it is accepted by Michelena (1954). Against this is the existence of the Aquitanian divine name ASTOILUNNO DEO, which, if it contains the word *asto,* would constitute evidence that *arsto* was a secondary form. See Agud and Tovar (1988-) for a survey of this question.

[34.] Bq **atal** 'segment, fragment, portion'

Bq a-**tal** 'limb', az-**tal** 'calf (of leg)', lower leg, heel' : Bur -l-**tAl**-ter 'foreleg (of quadruped)', 'shoulder (of horse)', 'arm (of human)' (B9la [33]) (B9lc PS1) (B9ld PS1)

Basque *atal* 'segment, fragment' does not have the meaning attributed to it by Bengtson. As for *aztal,* there is no reason to suppose this is related to *atal*; it looks like a derivative of *(h)atz* 'paw', but the second element is obscure: *atal* would not lose its initial vowel in this position.

[35.] Bq **atso** 'old woman'

Bq a-**tso** 'old woman' : PNC ***ʒʒʷVjV** 'female' : Bur -n-**co** 'paternal aunt' (B9la [135]) (B9lc PS1) (B93)

The Caucasian resemblance was pointed out by Bouda.

[36.] Bq **atzo** 'yesterday'

Bq **atso** 'yesterday' : Abkh **jacə**, Adyghe **dəRʷ-asă**, etc. 'yesterday' (C85 [24])

A minor error: the Basque word for 'yesterday' is *atzo,* while *atso* means 'old woman'. Many have tried to relate *atzo* to *atze* 'back', whose earlier form was almost certainly **atz,* but the final *-o* is inexplicable. Bouda (1948:§41) also points out the match with Abkhaz.

[37.] Bq **axuri** 'young lamb'

Bq a-**xuri** 'lamb' : Hattic **wa-zar-** 'ewe' (B9lb [33]) (B9lc PS1)

The Basque word is *asuri* ~ *azuri* ~ *a(t)xuri*. Some have seen this as a compound involving *zuri* 'white', and indeed a formation like **ardi-zuri (ardi* 'sheep') would account rather well for the phonology. Others have noted that this same word is used to call sheep, and proposed that the interjection might simply have been transferred to denoting young lambs.

[38.] Bq **azal** 'skin', 'bark'

Bq **azal** 'skin' : PAA ***c^{w}a** 'skin' (C85 [51])

There is a certain amount of evidence suggesting that the earliest form of the Basque word might have been **kazal*: some eastern varieties have *kaxal* 'bark' and also 'acorn cap', and Z has the apparent derivative *kazalda* 'dandruff'; moreover, a common word for 'fingernail' is *azkazal* (*hatz* 'finger'). Loss of initial **k-*, though rare, is apparently attested elsewhere (see under **astigar** 'maple'), and moreover an initial **k-* points indisputably to a loan word. The Caucasian form seems too short to offer a convincing match.

[39.] Bq **azeri** 'fox'

Bq **axeri** 'fox' (and Spanish **zorro**, of Basque origin) : Lezgian **ziru**, Ginukh **zeru**, Khwarshi **zaru**, Abaza **ser**, Andi **sor**, etc. 'fox' (C85 [72b])

The same or similar parallels with Caucasian forms have been pointed out many times before, including at least by Schuchardt (1913:113) (who rejects the comparison), by Uhlenbeck (1924), by Trombetti (1925:§114) and by Lafon (1948). At first glance, this looks like one of the more striking items in C's list, since the parallel, unusually, involves two consonants and a high degree of phonetic similarity. Unfortunately, it can be shown that the parallel is quite illusory.

Here *axeri* is the palatalized form of the more usual *azeri,* which has a widespread variant *azari.* This variation points strongly to an earlier **azeari*, which is supported by the Bizkaian forms *azegari* and *azagari* reported in Azkue's 1905 dictionary. Given the western propensity for inserting a *g* to separate vowels in hiatus, these forms in turn strongly suggest a lost intervocalic consonant, most likely an *n*. Michelena (1949a, 1949b) therefore proposes **azenari* as the original form of the word for 'fox', and notes with satisfaction that both *Azeari* and *Azenari* are widely attested as surnames in the medieval Basque Country, parallel to the widespread use of other animal names as surnames, such as *Otsoa,* from *otso* 'wolf' .

Michelena later (1956:§1; 1973:§119) goes further and proposes that the name of the animal is in fact *derived* from the personal name, which he traces to the Latin name ASINARIU or ASENARIU, which is well-attested in France and Spain, a conclusion which was apparently reached independently by the romanists Luchaire and Meyer-Lübke. In support of this, he cites the parallels of French *renard* 'fox', derived from the personal name *Reginhard*, the Bizkaian Basque *luki* 'fox', derived from the Latin personal name LUC1U, and the Alavese and Riojan Spanish *garcía* 'fox', derived from the common surname *García,* itself often thought to be of Basque origin.

The effect of all this is to destroy the claimed parallel with Basque *azeri.* That still leaves Spanish *zorro,* of course. This word has often been considered as being of Basque origin, in spite of the fact that it is completely unattested in Basque, but, in a comprehensive recent survey of this word, Corominas and Pascual (1980: under *zorra*) flatly reject a Basque origin on both phonological and distributional grounds: on the one hand, the word is first attested in Spanish with a *voiced* sibilant, making a Basque origin impossible; on the other, its first attestations are recorded in an area far from the Basque Country.

[40.] Bq **azkoin** 'badger'

Bq **azkonar** 'badger' (< ***harz**-konH-, where the first element is *hartz* 'bear') : PY ***χas** 'badger' (SR94 [11])
Bq harz-**koin** 'badger' (< *hartz* 'bear' + **-konH*) : PY ***kūñ** 'wolverine' (SR94 [303])
Bq haz-**koin** 'badger' (lit., 'bear-dog') : PW ****KUAN** 'dog' (BR94 [9])

The word for 'badger' is cited both as **hazkoin* and as **harzkoin*, but these forms do not exist: the word is *azkoin* ~ *azkoĩ* ~ *azkon(a)* ~ *azkuin* ~ *asko* in the dialects with the aspiration, save only for the variant *hazku* ~ *harzku,* found only in Zuberoan, the dialect which is known to have extended the aspiration to initial position in words which did not historically have it. Moreover, no such word as **koin* 'dog' is attested anywhere in Basque: this is a pure invention of the authors, and yet another impossible form with an initial voiceless plosive. Anyway, the dialect

variants point clearly to an original **azkone,* or possibly **arzkone* (Michelena 1949a:486; 1964b:193). Bengtson and Ruhlen are not the first to see the first element as representing *hartz* 'bear', but the second element is simply unidentifiable; see Agud and Tovar (1988-) for a survey of proposed etymologies for this word. See also **hartz** 'bear' below.

[41.] Bq **aztal** 'calf, heel: see under **atal**.

[42.] Bq **-ba** (in kinship terms)

Bq **ba** 'son', 'child' : Abkh a-**p**-ha 'daughter', etc. (C85 [64])
Bq **-pa** ~ **-ba** (alha-**ba** 'daughter', a-hiz-**pa** 'sister', ne-**ba** 'brother', etc.) : PY ***pu?-n** 'daughter', ***pu?-b** 'son' (SR94 [60])

No such Basque word as the putative **ba* exists. C has chosen to extract this morph from the words *alaba* 'daughter' and *ahizpa* 'sister (of a woman)' and to assign it an arbitrary meaning to suit his purposes. In fact, a morph *-ba* occurs in several other kinship terms: *arreba* 'sister (of a man)', *neba* 'brother (of a woman)', *osaba* 'uncle' (in places also 'father-in-law'), *izeba* 'aunt' (in places also 'mother-in-law'), *iloba* 'niece, nephew, grandchild'. There is little doubt that this morph must represent an ancient morpheme, but its range of occurrence is hardly consistent with the meaning of 'son' or 'child', any more than with the 'father' that the Basque nationalist Sabino de Arana wanted to assign to it. Interestingly, instead of *izeba,* western dialects have *izeko* for 'aunt', where *-ko* looks suspiciously like the familiar Basque relational suffix which forms words meaning, roughly, '(something) pertaining to', '(something) connected with'.

[43.] Bq **bai** 'yes'

Bq **bai** 'yes' : Bur A-**wa** 'yes' (B9lc [18])

[44.] Bq **barakurkuilo** 'snail'

Bq **barakuilu** 'snail' : PNC ***wHörλʷVɬV** 'snake, worm' (also **baIrčalu** 'snail' in a dialect of Lak) : Bur **γArkAs** ~ **γArqAs** 'lizard' (B9lb [35]) (B9lc [16]) (B9ld)

Basque *barakuilo is* no more than a doubtful *hapax* for the more widespread *barakurkailu* ~ *barakurkailo.* This is a compound whose second element is obviously of the same origin of the widespread Romance word represented by Spanish *caracol* 'snail', itself of doubtful origin, but certainly not Basque (see Corominas and Pascual (1980)). Michelena (1977: 332) identifies the first element as *bare 'slug'.* Remarkably, Bengtson (B9ld PS2) proposes to interpret the Romance word as a loan from Macro-Caucasian, but a Basque origin for the word is out of the question on phonological grounds.

[45.] Bq **bare** 'spleen', 'slug'

Bq **bare** 'spleen' : Sum **bar** 'liver' (B91c PS2 [34])

The Basque word means both 'spleen' and 'slug', and its apparent derivative *barauts* (*huts* 'bare, plain') means 'mucus, spittle'. (Cf. Castilian *baba* 'spittle, mucus' and its derivative *babosa* 'slug'.) This has led some to suppose that the earliest sense of the Basque word might have been something like 'mushy, mucky' or 'slime, slobber'; see the essay in Agud and Tovar (1988-).

[46.] Bq **bargo** 'young pig (3-6 months)'

Bq **bargo** 'small pig' : PNC ***wālrλ'ə** 'pig, hog' (B91c [74])

The Basque word is maddeningly similar to a whole set of western European words for 'pig', ranging from Romance *porca* 'sow' (of Indo-European ancestry) through Proto-Germanic **farhoz* 'pig' (of IE origin) and Spanish *verraco* 'boar' (of Latin origin) to Old English *bār* 'boar' (of unknown origin), and there have been the most strenuous attempts to drive it from some Indo-European source, so far unconvincingly.

[47.] Bq **barrabil** 'testicle'

Bq **barrabil** 'testicle' : PY ***bajbVl** 'kidney' : Sum **bir** 'kidney' (B91a [50])

The final element of the Basque word is the suffix *-bil* 'round', which is found in a number of other words *(gurpil* 'wheel', *ukabil* 'fist', etc.). The element *barra-* recurs in the words *barret* and *barrasaket* 'sterile' and *barratta* 'having only one testicle', suggesting that its central meaning is something like 'generation' or 'fertility'.

[48.] Bq **bartz** 'louse'

Bq **bartz** 'nit' : WCauc ***ć'a**, ECauc ***nęmʒV** (< ***welʒe**) (B94a [13]) (B94b [21])

Basque *bartz ~ p(h)artz ~ patz* 'louse, nit', with its widespread initial *p-*, does not look like a native word, and indeed Corominas and Pascual (1980) derive it unhesitatingly from the common western Romance element *bar- ~ var-* 'larva' (Portuguese *barro,* Occitan *vare,* etc.) See Agud and Tovar (1988-) for further Romance parallels and a possible Latin source.

[49.] Bq **baso** 'woods'

Bq **baso** 'woods' : PAT ***maza** 'pine tree', 'spruce', 'conifer', Adyghe-Kabardian **măz(ə)** 'woods' (C85 [79])
Bq **baso** 'forest' : PNC ***wɨce** 'mountain' (B91b [26]) (B91c [15])

The word *baso* is today the usual term for 'woods, forest' in the west of the Basque Country, while eastern varieties have *oihan.* There is evidence from compounds, however, in all parts of the country, suggesting that the earlier meaning of *baso* was merely 'wilderness', 'uncultivated or unsettled land'. Moreover, in the predominantly deciduous Basque forests, dominated by ash, oak, beech, and birch, conifers were not (until recently) an especially prominent feature of the landscape.

Note that C and B do not agree in their choice of proposed Caucasian cognates. These parallels have been pointed out several times before, including by Uhlenbeck (1923:24) and by Bouda (1948:§62).

[50.] Spanish **becerro** 'bullock'

Sp **becerro** 'calf, young bull' : PNC ***wělš(w)ē ~ *bělš(w)ē** '(bull) calf' : Bur **b(uš)ōšo** (B91b [32])

Presumably B means to suggest that this Spanish word is of Basque origin, but such an etymology is categorically rejected by Corominas and Pascual (1980).

[51.] Bq **begi** 'eye'

Bq **begi** 'eye' (? be-**gi**) (< ***bergi** or ***berγi**) : PNC ***ʔwilhi** 'eye' (WCauc ***b-l'a**) : Bur **-l-či(n)** 'eye' : Sum **igi ~ ibi** 'eye' (B91a [3]) (B91c PS1') (B91d) (B94a [9]) (B94b [10])

The universal Basque word for 'eye' is *begi*; this word is attested early and occurs in a large number of seemingly ancient compounds. It is not clear whether B wants to remove the "body-part prefix" *be-*, as he so often does, since he seems to need the *b-* this time to match the PNC *w*. B obtains his putative **berγi* solely from the Z form *berphuru* 'eyebrow' (better, *berp(h)ürü*), but this idiosyncratic form is correctly described by Lhande (1926) as a "corruption" of the otherwise universal *bepuru*, a compound of *begi* + *buru* 'head'. B wants this form in order to obtain a segment to match the *l* in the Caucasian form, but he can't have it: the Z variant results merely from an anticipatory assimilation, and the matchup is feeble.

[52.] Bq **beginini(ko)** 'pupil (of the eye')

Bq begi-**nini**-(ko) 'pupil of the eye' : PNC ***nănV** 'pupil of the eye'; 'doll' : Bur **nAna** 'eveball'; 'uncle' (B9la [4]) (B91d PS2) (R94b)

Basque *nini* means both 'child' and 'pupil' (and also 'doll, toy'). Bengtson (1991d:168) notes the obvious resemblance to Spanish *niño* 'child' and *niña* 'pupil', but amazingly concludes that the Spanish word is borrowed. This is hardly likely: Basque *nini* cannot possibly be ancient and is surely an "expressive" formation if it is not merely a loan from Romance.

[53.] Bq **behatz** 'toe' : see **hatz** 'finger'.

[54.] Bq **behazun** 'bile'

Bq be-**azun** 'bile, gall' : PAA ***za**, PND ***c̄wümh̄i** (C85 [53])

Bq be-ha-**z(t)un** 'bile, gall' : PNC *** c̄wümhi** 'bile, gall, anger' : Bur **čhęmI-liŋ** 'bitter, poisonous, poison' : PY ***sen** (B9lb [10]) (B9lc PS1) (SR94 [154])

Yet again C and B are removing their "body-part prefix", though the notion of 'body part' would hardly seem to stretch this far. This word shows a range of local variants: *be(h)a(t)zun*, *behazun(e)*, *be(r)aztun*, and *beazuma*, the second of which also means 'affliction' and the third of which also means 'gall bladder'. The eastern dialects have *bedamin*, with a different ending. Since the form *behazane* seems to involve the derivational suffix *-une* discussed below under **(g)une**, and since the *-tun* of *beaztun* also looks like the common derivational suffix *-dun* 'having' (devoiced to *-tun* after a sibilant), it rather looks as if the stem of all these forms is something like **be(h)a(t)z-* or **beCa(t)z-*, rather than the putative **-azun* demanded by C and B. Agud and Tovar (1988-) make a vigorous case for deriving the Basque word from *begi* 'eye'; this is too long to reproduce here.

Similar parallels are drawn by Bouda (1952: § 29).

[55.] Bq **beko ~ moko** 'forehead'; 'beak, extremity'

Bq **bekho** 'face, beak' : PNC ***běk'wV** 'mouth' : Bur **bu:k** 'throat, neck' (B93)

Bq **mokho** 'beak, front, face', **mokhoz-mokho** 'face to face', **mik**-usi 'see, perceive' : PNC ***mīqqV** 'moustache' : Bur **mōq-Iṣ̌** 'cheek' (B9la [5]) (B9lc [102])

It is generally believed by vasconists that *moko* and the largely archaic *beko* are variants of a single word. There is good reason to believe that the Basque word is of Romance origin; see the long discussion in Corominas and Pascual (1980) under *mogote* (the usual Romance variation **boc- ~ *buec-* would account nicely for the Basque variants). On the other hand, Meyer-Lübke (1935) identifies a late Latin BECCU 'beak, bill', widely attested throughout western Romance (including Italian), and Basque *beko* may simply be a borrowing of this. As for Bengtson's alleged "*mikusī*" 'see', this does not exist. The word for 'see' is *ikusi*, and I presume what Bengtson is trying to cite is the phrase *ikusi-mikusi* (or *ikusi-makusi*), the name of a children's game resembling "I spy". This is inattentive, to say the least.

[56.] Bq **bekoki** 'forehead'

Bq be-**koki** 'forehead, crown, nape' : Bur **-kAk** 'crown, top of the head' (B9lb [2]) (B91c PS1) (B93)

Agud and Tovar (1988-) review several plausible-looking etymologies for *bekoki*. The most obvious one derives it from *beko* (above), which is well attested in the sense of 'forehead' in the archaic language, plus the common noun-forming suffix *-ki*. Several scholars have preferred to see this word as a derivative of *begi* 'eye', and it occurs to me that a formation **begi-goi-ki (goi* 'high part') would straightforwardly yield **bekoiki*, whose dissimilation to *bekoki* would be in no way unusual in Basque.

[57.] Bq **belarri** 'ear'

Bq be-**larri** 'ear' (*be-* = body part prefix) : PAA ***ÍV**, PND ***lěrħIV** (C85 [48])
Bq be-**larri** ~ be-**harri** 'ear' : PNC ***lě(r)ħIV** 'ear' (WCauc ***l'V**) (B9la [8]) (B9lc [78]) (B93) (B94a [21]) (B94b [32])

There are several problems here. First, an intervocalic *l* in pre-Basque does not survive in modern Basque, but is regularly converted into *r* (Michelena 1977: § 16.2). Hence ancient Basque could not have had an *l* in this word. Second, *belarri* is only the western Basque form for 'ear'; eastern varieties have *beharri* or *begarri*. This induces Michelena (1977:339) to reconstruct **berarri* as the original form of this word, with different dissimilations occurring in the various dialects. Such dissimilation of intervocalic *r* before a following *rr* is a regular feature of the phonological history of Basque (cf. **bular** 'breast' and **elur** 'snow' below). Third, C and B's confident removal of the morph *be-* as a "body-part prefix" is again without justification.

[58.] Bq **belaun** 'knee'

Bq be-**la**-un 'knee' (*be-* = body part prefix; *-un* = suffix) : PAA ***xa** 'leg'; PND ***ǩělħV** 'leg' (C85 [47])
Bq be-**lhaun,** be-**lhaur-** 'knee ' : PNC ***łwilV** 'elbow' (B93)

Once again C is indulging in some outrageous arbitrary segmentation of the Basque word. As usual, he removes his "body-part prefix" *be-,* and then he removes a "suffix" *-un* for no better reason than that Basque *lagun* 'companion' also ends in *-un*, and he wants to remove that *-un* too. B follows C in removing *be-,* though he doesn't remove the ending; however, he adduces the combining form *belaur-* in order to get a liquid into final position. But this is unjustified: ancient Basque stems in *-n* frequently show a combining form in *-r,* for unknown reasons. Compare Bengtson's treatment of **egun** 'day' below, in which he ignores its combining form *egur-* because there it is the *n* he wants for his match. But he can't have it both ways: he can't pick and choose among alternants at will in order to find matchups.

Bouda (1949:§85) makes very similar observations about *belaun* and North Caucasian words for 'leg'.

[59.] Bq **beltz** 'black'

Bq **beltz** 'black' : PNC ***ʕʷɔ̄lɔ̄č'V** 'black' (B9lb [60])

There are question marks here. To begin with, Aquitanian exhibits an element *Beles-*, which can plausibly be identified with *beltz*, suggesting that this form is a contraction of earlier **beletz*. Further, there is evidence that the color term is itself a derivative of an element **bel(e)-*, in the form of such words as *bele* 'raven, crow' and *harbel* 'slate' (? *harri* 'stone' + **bel* 'dark'). If this thinking is right, the matchup is destroyed.

[60.] Bq **berezi** 'separate'

Bq **ber**-hezi 'separate, distinguish' : Bur **bAr-** 'thresh' : Sum **bar** 'split, open' (B9lc PS2 [11])

Basque *berezi* 'separate, distinguish' is analyzed as containing an alleged root **ber-* of the same meaning; in fact, it derives from *berez* 'by itself', the instrumental case of the common element *ber-* 'self, same'.

[61.] Bq ****beri**

Bq **beri** 'this same' : Abkh a-**bri** 'this' (= *a-bari*), etc. (C85 [17])

The putative Basque **beri* does not in fact exist. There is a stem *ber-* 'self, same', which serves as the base of several derivatives, but the one which means 'this same' is *berau,* in which the second element is the proximal demonstrative *hau* 'this'. The corresponding form for the distal demonstrative is *bera,* which means '(he) himself', '(she) herself'; western dialects have recently generalized this form into an ordinary third-person pronoun 'he', 'she'. The stem *ber-* never means 'this', and hence the comparison fails on semantic grounds.

Dumézil (1933:140) and Lafon (1951:§15) compare Basque *ber-* with precisely the same North Caucasian forms.

[62.] Bq **bero** 'hot'

Bq **bero** 'hot, warm' : Bur **bAbAr**-um 'hot, warm' (B9lc [10])

Most vasconists are satisfied that Basque *bero* is a native word, though some have been attracted by the Gaulish Celtic **borvo-* 'hot', cognate with Breton *berv* ~ *bero* 'cooked' and other Celtic forms. Agud and Tovar (1988-) present an impressive catalogue of similar words, most of them previously compared with Basque *bero,* ranging from ancient Egyptian *brbr* 'boil' through Georgian *birbili* 'burning' to Maori *wera* 'hot', and including earlier citations of the Burushaski word. The problem is that an onomatopoeic item of the approximate form "*berber*" seems to be almost universally used to represent boiling, and often acquires transferred meanings like 'hot', 'burn' or 'cook'; consequently, items with such forms and meanings are likely to prove very treacherous in comparative work.

[63.] Bq **berri** 'new'

Bq **berri** 'new' : Coptic **bēri, brre** 'new, young' (B94b)

Bengtson remarkably sees the Basque word as a loan from ancient Egyptian. In fact, many previous linguists have pointed to the Basque-Coptic resemblance and tried to interpret it in one way or another, but there seems no reason to suspect anything beyond a chance resemblance.

[64.] Bq **beso** 'arm'

Bq **beso** 'arm': PND ***bač'V** 'paw' (C85 [46])

This resemblance is noted by Bouda (1948:§62). Note that C, in contrast to his usual practice, does not remove the "body-part prefix'' *be-* in this case because this time he needs the *b-* to get his match. But he can't have it both ways: he can't arbitrarily keep or remove the morph *be-* to suit his convenience.

[65.] Bq **bete** 'full'

Bq **bethe** 'full' : PNC ***bVtV** 'solid, strong' : Bur **bu:t** 'much, many, numerous, very' : PY ***bid-** 'strong' (B94b [30]) (SR94 [261])

[66.] Bq **bi** 'two'

Bq **bi** 'two' (< **ɢʷi*, assimilated to *bat* 'one') : PNC ***(t)q'ʷiI*** 'two' (B9ld) (B94a [2]) (B94b [3])

B's **ɢʷi* represents nothing more than speculative special pleading. Eastern Basque has a longer form *biga,* and it is known that the numeral for 'two' was anciently postposed in Basque (as it still is in B today), and Michelena (1977: 413) therefore concludes that *bi* results from a reduction of *biga* to **bi(a)* in this postposed position, destroying the proposed matchup.

[67.] Bq **bide** 'road, way'

Bq **bide** 'road' : Berber a-**brid,** ta-**brida** 'road' (B94b)

Bengtson sees the Basque word as a loan from Berber, which appears to be no more plausible phonologically than the oft-cited idea that the word is a loan from Latin VIA 'road'.

[68.] Bq **biga** 'two-year-old heifer'

Bq **biga** 'heifer' : PNC ***hp̄äX̌'ʷE*** 'small stock' (B9lb [31]) (B9lc [6], [73])

Basque *biga* is also widely attested as *bigae*, showing the loss of a consonant, and Michelena (1977:145) confidently reconstructs **bigana*, destroying the match. This word, which means 'two-year-old heifer', is either a derivative of *biga* 'two' or, more likely, a loan from a Latin *BIMANA 'two years old' with nasal dissimilation (Michelena 1974: 201); descendants of this Latin word are widely attested in western Romance in the general sense of 'heifer'.

[69.] Bq **bihar** 'tomorrow'

Bq **biha**-r ~ **biga**-r 'tomorrow' : PNC ***p̄əkV*** 'dawn' : PY ****pʌk-*** 'morning' (B9lc [5]) (SR94 [168])

As usual, Basque *h is* not etymological, and western *bigar* is probably only a secondary variant showing the common western insertion of *g* to separate vowels in hiatus; the attempted matchup therefore fails. There have been strenuous efforts to derive the word from *bi ~ biga* 'two', and to reconstruct the sense as 'second day'; these efforts have not so far won general acceptance, but, if successful, would destroy the matchup.

[70.] Bq **bihotz** 'heart'

Bq bi-**hotz** 'heart'; 'love, courage, etc.' : Bur **-As** 'heart, mind', **-As**-ki 'remembrance, longing, desire' : Sum **úš** 'blood'; 'love, mercy, understanding, etc.' (B91a [48]) (B9lc PS1) (B9ld) (B93) (B94b [11])

Basque *bihotz* appears to be represented in several Aquitanian names, such as *Bihoxus* (it is known that Aquitanian *x* represents an affricate). It is difficult to see that the very vague resemblances to Burushaski and Sumerian constitute any improvement on the colorful collection of fanciful etymologies catalogued in Agud and Tovar (1988-), and the appeal to transparently secondary senses like 'love' does nothing to improve a rather feeble comparison.

[71.] Bq **bildu** 'collect, gather'

Bq **bil-** 'gather, unite, amass' : Bur **bil ~ bir-bir** 'full' (B9la [108]) (B9lc PS1')

In spite of strenuous attempts to derive Basque *bildu* from some Romance source such as a Latin *PILARE 'pile up', it seems more likely that the stem *bil-* is native. It does not, however, mean 'fill' or 'full'.

[72.] Bq **birika** 'lung'

Bq bi-**rika** 'lung' : PNC (East) ***jerk'ʷi** 'heart' (B9lc [92]) (B93) (B94a [10])

Unfortunately for Bengtson, the word for 'lung' in the earliest Basque texts is *biri*, a form which is still current in parts of the country today. It appears that the morph *-ka* is a late accretion to the word, thus destroying the proposed match.

[73.] Bq **bizar** 'beard'

Bq **bizar** 'beard' : PND ***moć'orV** (C85 [52])

Note that the "body-part prefix" is *not* removed in this case, since the *b-* is needed for the comparison. Bouda (1948: §106) compare Basque *bizar* with a range of North and South Caucasian forms, including the Nakh-Daghestan words. Other scholars have variously compared the Basque word with different Caucasian forms and with words from Berber, Paleo-Siberian, Altaic, Romance, and other sources; see Agud and Tovar (1988-) for a catalogue.

[74.] Bq **bizi** 'alive'

Bq **bizi** 'alive, living' : Abkh a-**bza** (PAT ***bəza**) (C85 [33])
Bq bi-**zi** 'life, alive, live (v.)' : PNC ***šiʔʷV** 'soul, breath' : Hurrian **šeγ-iri** 'alive' : Urartean **šu/oh**-ori 'alive' (B9lb [14]) (B9lc [61]) (B93)

Basque *bizi* is an adjective ('alive') or a noun ('life'); it is not a verb (the verb 'live' is the compound *bizi izan*). It is clear from the evidence of such derivatives as *biztu* ~ *piztu* 'ignite, animate, revive', stem *bitz-* ~ *pitz-*, that *bizi* is to be segmented as *biz-i*. Michelena (1950a:463) observes a resemblance between this word and Georgian *-ghvidz-* 'awake, awaken', and notes that the phonological matchups are impeccable, in terms of the Basque-Georgian correspondences recognized by Lafon (1948); I do not know whether C would also accept such a match. Michelena also finds the semantic connection acceptable, but he nonetheless notes with approval the suggestion of Schuchardt (1923:§37) that Basque *bizi* is derived from some Romance (possibly Occitan) development of Latin VIVU 'alive', though the phonological development does not appear to be regular.

Bouda (1948:§42) also compares Basque *bizi* to the Abkhaz form cited by C and to other Caucasian forms.

Bengtson, as so often, differs from Čirikba in his choice of Caucasian cognate; his removal of the alleged "body-part prefix" *bi-* from *bizi* strikes me as simply outrageous.

[75.] Bq **bizkar** 'back', 'height in mountains'

Bq bi-**zkar** 'back' (of body) : Abkh a-**zkʷa** (C85 [49])
Bq bi-**zka**-r 'back' : Abkh a-**zkʷa** 'back' : Bur **-sqa** 'on one's back' : PY ***suga** 'back' (adv.) (B9lc [85]) (B93) (R94b) (SR94 [10])

This time Čirikba and Bengtson agree in their proposed matchups, but once again they are arbitrarily spiriting away that portion of the word *bizkar* which they find inconvenient. Anyway, the positing of a Basque "morpheme" **-zkar* is preposterous: no ancient Basque stem can begin with any consonant cluster at all, let alone with a sibilant-plosive cluster, which Basque speakers find every bit as difficult as Spanish speakers; initial clusters in Latin loan words into Basque were, without exception, reduced in one way or another.

The word *bizkar* is clearly ancient. It is universally thought to be related to the name of the Basque province *Bizkaia*, and Michelena (1954:455) notes the existence in the ancient Iberian texts of a word *biscar*, which he is inclined to identify with Basque *bizkar* (presumably as a loan word in one direction or the other — Michelena flatly rejects a close genetic link between Basque and Iberian).

[76.] Bq **bortz** 'five'

Bq **bortz** ~ **bost** 'five; hand' : PNC ***bɔ̄rcV** 'paw; gammon (hog's thigh)' : Bur **bAc-in** 'thigh' (B9la [30]) (B9lc [7], [38]) (B94b [25])

There is no doubt that eastern *bortz* is the more conservative form, with western *bost* being an innovation. Bengtson's claim that Basque *bortz* ~ *bost* means 'hand' is a fantasy: the word means nothing but 'five'. Azkue records the derivative *bosteko* 'hand', which he describes as "trivial" — in other words, as jocular or slang, as in the American English phrase "Gimme five!". Bengtson's semantic match is unconvincing in the extreme.

[77.] Bq **bular** 'breast, chest'

Bq **bul(h)a-**r ~ **bura-**r ~ **buda-**r 'breast, bosom' : Sum **u-bur** 'breast' (B9lc PS2 [20])

It is far from clear whether the widespread (and early-attested) *bular*, or the severely localized *burar*, is the older form, with the weight of evidence perhaps favoring *burar* (Michelena 1977:315); dissimilation of *burar* to *bular* would be quite normal in Basque (compare **belarri** 'ear' above). Agud and Tovar (1988-) provide a catalogue of attempted etymologies and cognates from Caucasian and other sources; noteworthy is Corominas's attempt (1969: 174) to relate the word to the Germanic item which appears in English as *breast*. Note that B does not remove his "body-part prefix" here.

[78.] Bq **buru** 'head'

Bq **buru** 'head', be-**p(h)uru** 'eyebrow, eyelash, eyelid' : (Caucasian) Udi **for** / **fufru-**x 'hair', Kurin **firi** 'mane' : Bur **buur** '(a single) hair', **bupuur** 'down, fine wool', **-l-puur** 'eyelash', **-Aš-puur** 'mane', **buuri** 'crest, peak' (B9la [13]) (B9lc [149])

There is a good deal of confusion here. First, Basque *buru* means only 'head'; it never has any meaning related to 'hair', or 'mane', or 'eyelash', or anything of the sort, though it does frequently have extended meanings along the lines of 'chief part of' or 'highest part of'. Second, the word *bepuru*, a transparent and phonologically regular compound of *begi* 'eye' and *buru,* means only 'eyebrow'; it does NOT mean 'eyelash' or 'eyelid' (these last two are *betile* and *betazal,* both compounds of *begi,* with *ile* 'hair' and *azal* 'skin', respectively. Bengtson's rather labored efforts at making his matchups go through are completely unconvincing, and note that he both removes and fails to remove his "body-part prefix", according to need.

[79.] Bq **busti** 'moist, wet'

Bq **busti** 'moisten' : PAT ***pəsə** (C85 [28])
Bq **busti** 'wet' : PNC ***wɨ̱ć^{w}V** 'wet' (B9lc [17], [56])

Basque has an adjective *busti* 'wet, moist' (with a severely localized variant *musti*), from which has been derived a verb *busti* 'moisten': the form shows unmistakably that the adjective was earlier. All vasconists and romanists are satisfied that the word is a loan from Romance, possibly directly from Latin MUSTEU 'fresh, moist', as preferred by Michelena (1971a:157), or else from Occitan *musti* 'moist', from a late Latin derivative *MUSTIDU (widely attested elsewhere in Romance), as preferred by Schuchardt, Corominas, and Tovar; this etymology makes the word cognate with English *moist*. (There are abundant parallels for the sporadic development of Romance *m-* into Basque

b-: B *bolu* 'mill' < *MOLINU (this is the first element in the famous Basque surname *Bolívar*); Z *bedezĩ 'doctor'* < French *médecin*); Z *benüsér* 'carpenter, cabinetmaker' < French *menuisier;* and others.)

[80.] Bq **d-** (verbal prefix)

Bq **d-** (3rd sg. marker) : Abkh **d-** (same, animate) (C85 [6])

This resemblance is noted by Dumézil (1933: 138) and by Lafon (1952a:§26); Lafon in fact goes further, citing a number of third-person pronouns and demonstratives in *d-* in North Caucasian languages.

But there are several problems here. First, Basque *d-* does not mark "third singular", but third-person absolutive (singular and plural). Second, it does so only in a restricted subset of forms — in the modern language, in present-tense forms, though sixteenth-century texts show a slightly broader distribution. Other verb forms have different prefixes for the third person: Bizkaian Ø- ~ common *z-* in the past tense (it is widely thought among vasconists that Ø- is probably the original prefix), *l-* in the so-called 'hypothetic' forms, and *b-* in the jussive.

This array contrasts puzzlingly with the uniform marking of first and second persons throughout the paradigm by morphs that correlate strongly in form with the independent personal pronouns. Several opinions have been expressed in the literature. Jacobsen (1975:18-21) regards *d-* as the "true" third-person prefix, deriving it from a lost pronoun, and dismisses the others. Schuchardt (1923) also regards *d-* as the genuine third-person marker (derived from a lost demonstrative), and tries to derive the other prefixes from incorporated lexical items with modal or aspectual meanings. Lafon vacillates on this issue. In his book (1944:387 ff.), he reluctantly endorses Schuchardt's view; in a later paper (1952a:82-83), he is inclined to posit a series of lost demonstratives as sources for the prefixes; in a late monograph, however (1973:88), he once again returns to endorsing Schuchardt's position.

In Trask (1977:204-205), I propose that this whole approach is misguided. Each of the four prefixes is clearly a simultaneous exponent of third person and of a tense or mood category. It therefore seems eminently plausible to assume that these prefixes were originally markers only of tense or mood and that they came to be lost when a first- or second-person agreement marker was prefixed to the verb form. In the third person, in which there was no prefix (and the third person is marked by zero in most other circumstances in the Basque verbal paradigm), the prefixes remained, yielding the observed distribution, without the necessity of postulating any ghostly pronouns or demonstratives. Among its other virtues, this analysis has the advantage that it explains why all the "third-person" prefixes appear in finite forms in which there is no relevant third-person argument, but in which, for one reason or another, there happens to be no first- or second person prefix — a fact which sent some of the earlier analysts into conniptions.

Whether or not the reader finds this analysis convincing, it is clearly too much for C merely to take Basque *d-* as a third-person singular agreement marker.

[81.] Bq **ebaki** 'cut

Bq **ebaki** ~ **ebagi** 'cut' : PY ***pak** 'cut' (SR94 [59])

Basque *ebaki* (B *ebagi*) is one of a sizeable number of verbs exhibiting a final morph *-ki* (B *-gi*). There is abundant evidence that this morph was anciently a suffix which changed the valency of the verb-stem to which it was attached (see, for example, Lafon (1944:II:14-18), Trask (1981, forthcoming *a*)). There is a second such morph in Basque, *-tsi,* and it is noteworthy that the verb *ebatsi* 'rob, steal' appears to show the same stem with this other suffix. In all probability, the earliest form of the root was **-ba-,* thus destroying the match (see Michelena 1977:231).

[82.] Bq **egin** 'make, do'

Bq **egi(n)** 'make, do' : Sum **ak** 'make, do' (B9lc PS2 [38])

Today the root of the Basque verb is usually *-gi-*, but there is clear evidence that it was originally **-gin-* (Trask 1990), and the resemblance to Sumerian is too faint to take seriously.

[83.] Bq **egun** 'day', **eguzki** ~ **eki** 'sun'

Bq **eguzki** 'sun', **ego** 'south', **egun** 'day' : Adyghe tə-**Ră**, Kabardian də-**Ră**, Ubykh d(ə)-**Ră**, nd-**Ră**, PND ***wiraGGV**, all 'sun' (C85 [86])

Bq **ekhi** 'sun' : PDC ****jakʰ-** 'sun' (B9la [67])

Bq **igu**-zki 'sun, day', **egu-** 'day' : PNC ***ʔʷɬGGɬ-nV** 'day' : PY ***xiGa** 'sun' : Sum **ug** 'sun, heat, light, day' : PDC ****HiGu-(N)** 'sun, day' (B9la [68])

Bq e-**gun** 'day, daylight' : PNC ***Gʷĕm-**tV 'day, 24 hours ': Bur **goon** ~ **gUn** 'dawn', **gUn-c** ~ **guun-c** 'day' : PY ***gəʔn** 'bright, light' : Sum **gun** 'bright' : PDC ****Gun-** ~ ****Gʷen-** 'day' (B9la [69]) (B9lc [110], PS1, PS1') (B9ld PS1) (R94b)

Bq e-**gun** 'day' : PY ***χōŋ** 'day(time)' (SR94 [62])

These entries exhibit monumental confusion. Basque *egun* is the universal word for 'day'; like other *n*-final stems, it shows a combining form *egu-*, as in *eguerdi* 'noon' (*erdi* 'middle'), or occasionally *egur-*, as in *eguraldi* 'weather' (*aldi* 'occasion'). The two forms of the word for 'sun' are both derived from *egun* plus the noun-forming suffix *-(z)ki*. The development of **egu-ki* to eastern *eki* is absolutely regular (compare *ekaitz* 'storm', from *egun* plus *gaitz* 'bad'), while *iguzki* is a variant of *egazki* showing raising of the first vowel before a high vowel (compare *ikusi* 'see', from earlier (attested) *ekusi*).

Somehow Bengtson manages to derive *egun* 'day' from a putative Dene-Caucasian root ****Gun**, while on the same page deriving its combining form *egu-* and its compound *eguzki* from a *different* D-C root ****HiGu(N)**, and its compound *ek(h)i* 'sun' from yet a *third* D-C root ** **jakʰ**-! There could hardly be a more sobering reminder of the dangers of trying to do comparative linguistics on the basis of mere resemblances in form and meaning.

Outside the obvious derivative *eguzki* ~ *eki*, there is no evidence in Basque for relating *egun* to the meaning 'sun', nor is there any evidence for relating Basque *hego* 'south' to *egun*. In fact, the best evidence we have is that *hego* originally meant 'south wind' (a meaning which it still has), and that the common compound *hegoalde* 'south' (*alde* 'side, region') is probably the source of the meaning 'south' (three of the four Basque compass points have names related to, and probably derived from, names of winds, and such naming seems to be a common practice in mountainous regions).

See also **belaun** 'knee' above.

[84.] Bq **egur** 'firewood'

Bq e-**gur** 'firewood' : PNC ***gə̄rV** 'stick, pole', etc. (B93)

[85.] Bq **ehiza** 'hunting'

Bq **eiza** 'hunt(ing)' : Abkh **a-šʷa-**, Ubykh **śʷa-ḱ'ă** etc. (C85 [81])

Basque *eiza* is the western form of the eastern *ihize*, *ihizi*. The existence in the eastern dialects of nasalized vowels, as in Zuberoan *ĩhĩze*, shows clearly that an intervocalic *n* has been lost, and that the word must have originally been **inizV* or **enizV* (neither the first nor the last vowel can be reconstructed with certainty); see Michelena (1977:115, 129). As so often, the former presence of this nasal has grave consequences for the proposed matchups.

Bouda (1948: § 8) notes similar parallels.

[86.] Bq **ekei** 'material'

Bq (Z) e-**khei** 'material' : PNC ***q'ajV** 'thing' (B9lc [106])

The usual Basque word for 'material' is *gai* ~ *gei*; the Zuberoan form is merely a compound of this with *egin* 'make' (combining form *egi-*); the development of **egi-gei* to *ek(h)ei* is phonologically impeccable.

[87.] Bq **elur** 'snow'

Bq e-**lhu**-r 'snow' : PNC ***ʎʷiu̯V** 'snow' (B9lc [67]) (B9ld) (B93)

Here as elsewhere, Bengtson appears to believe that there is something distinctive about the Basque aspiration; there is not. A severe difficulty with the proposed matchup (quite apart from the unexplained final *-r* in the Basque word) is that the variants *erur* and *eur* are attested early, and that *edur* (from **erur) is* the general B form today. This leads most vasconists to suppose that **erur* is the original form, with the same dissimilation of *-r- to -l-* that appears in **belarri** 'ear' above (the final *-r* in *elur* is trilled: definite *elurra* 'the snow').

[88.] Bq **emakume** 'woman'

Bq ema-**k(h)ume** 'woman' : PNC ***qʷänV** 'woman' : Bur **quma** ~ **kuma** 'concubine' : PY ***qVm-** 'woman' : Sum **géme** 'maiden, slave' (B9lb [40]) (B91c [100]) (SR94 [305])

Basque *emakume* (Bengtson's **emakhume* does not exist and would be impossible), which incidentally is not attested in our earliest texts, is a transparent compound of *eme* 'female' (combining form *ema-*) and *-(k)ume* 'child, offspring' (Michelena 1950a:468); possibly the original meaning was 'girl', the sense which this word has today in the Salacenco dialect. This destroys the matchup. Moreover, *eme* itself is in all likelihood a loan from Romance: cf. Occitan *heme,* Bearnais *hemne* 'female', from Latin FEMINA.

[89.] Bq **emazte** 'wife'

Bq ema-**zte** 'woman, wife' : PNC ***cVdV** ~ ***ćVdV** 'female' : Hurr **aš̄ti** 'woman, wife' (B9lb [39]) (B91d PS1)

It can hardly be doubted that the first element of this word is again *eme* 'female', discussed in the preceding entry. The second element is obscure, though several have seen it as *gazte* 'young': a formation **ema-gazte* would be phonologically impeccable, though the semantics would require some justification. In any case, the matchup is destroyed.

[90.] Bq **eme** 'sweet'

Bq **eme** 'sweet' : PNC ***HvnēʔʷV** 'sweet' (B9lc PS1)

[91.] Bq **entzun** 'hear, listen to'

Bq **entzu**-n 'hear, listen' : PNC ***ʔamc'V-** 'know, see' : PY ***ʔVt-** 'know' : Hurr **anz**-an-uy- 'declare (?)' : Sum (ni-)**zu** 'know' (B9lb [64]) (B9lc PS2 [25])

The Basque verb *entzun* has a very odd shape: since *e-* must here be the prefix invariably found in ancient verbs, that leaves us with an implausible-looking root of the form *-ntzu-*. Michelena (1977:114) reconstructs **enezun,* with an ancient syncope, a reconstruction which does nothing to improve the already dubious match with Caucasian and Yeniseian.

[92.] Bq **erdi** 'middle', 'half'

Bq **erdi** 'half, middle' : Bur **ālto** 'two' : PY ***ʔa(ʔ)l** 'half' (SR94 [120])

The catalogue of proposed cognates for Basque *erdi* provided by Agud and Tovar (1988-) includes such gems as Sanskrit *ardha-* 'middle, part', Lithuanian *ardyti* 'separate', Kartvelian *rt* 'divide', Georgian *kerʒ* 'middle' Laz *kuerdi* 'middle, half', and western Dravidian *eradu* 'two', among others. It is hard to see the Burushaski and Yeniseian proposals as improvements.

[93.] Bq **eri** 'finger'

Bq **erhi** ~ **erri** 'finger, toe, finger's breadth (measure)' : PNC ***rem**-k̆'V 'cubit', *HV-**rVm**-k̆V 'armful, bosom' : Bur **-riiŋ** ~ **-reen** 'hand' (B9la [26])

The Basque word is *eri,* or *erhi* in the aspirating dialects, but never **erri.* Bengtson's invocation of the transparently secondary sense of 'finger's breadth' strikes me as a rather desperate attempt to inject some element of plausibility into what is a *very* unconvincing parallel.

[94.] Bq **eri** 'sick'

Bq **eri** 'sick, sickness' : Bur **her-** 'weep' : Sum **ér, ir$_6$** 'weep' (B9lc PS2 [26])

[95.] Bq **ero** 'crazy', 'stupid'

Bq **ero** 'stupid, fatuous' : PY ***ʔaru-** 'sly/deceive' (SR94 [239])

It is very difficult to see the point of this semantically curious comparison. In any case, eastern varieties of Basque have a lateral in the word (R Sal *ello,* Z *elho*), which strongly suggests an original **elo.*

[96.] Bq **esku** 'hand'

Bq **esku** ~ **eski** 'hand' : Bur **hiisk** ~ **hesk** 'wrist; back of the hand' (B9la [29]) (B9ld) (B93) (B94b [24])

The Basque word is *esku,* never **eski.* What Bengtson is trying to cite is the Z form, in which the *u* has been fronted to give *eskü,* and the addition of the article produces unrounding of *ü*: *eskia.*

[97.] Bq **eskubarne** 'palm (of the hand)'

Bq (esku)-**bar**-ne ~ (esku)-**barr**-en 'palm' : PNC ***p̄^{w}ar-k̄^{w}V** ~ ***bar-k̄^{w}V** palm of the hand' : PDC ****bar-** 'palm' (B9la [36])

Basque *eskubarne* 'palm of the hand' is fantastically analyzed as containing an alleged root **bar-* 'palm'; in fact, the word is merely a compound of *esku* 'hand' and the common *barne* ~ *barren* 'interior'; such compounds are very frequently used in forming body-part names in Basque.

[98.] Bq **esne** 'milk'

Bq e-**sne** 'milk' : PNC ***šV̲nHV** 'milk' : PY ***de(ʔ)n** 'milk/nipple' (B9 lc [65]) (B93) (SR94 [165])

[99.] Bq **etorri** 'come'

Bq **e-thorri** 'come, attend, proceed' : PNC ***ʔV-t'Vr-** 'run' : Sum **tur$_5$** 'enter, bring in' (B9lc [20], PS2 [50])

Basque *etorri* just means 'come', never 'attend' or 'proceed'.

[100.] Bq **etxe** 'house'

Bq e-**tze** ~ e-**txe** ~ i-**txe** 'house', e-**txo**-la 'hut, cabin' : PNC ***c'VrHV** 'house' : Sum **éš**, 'house' (B9lc PS2 [29]) (B9lb [49])

Though *etxe* is by far the most widespread form of the word for 'house', its *tx* marks it unmistakably as an "affective" form, and most vasconists believe that the original form was the (attested) *etse*, not cited by Bengtson. The word *etxola* is no more than a derivative of *etxe* (Michelena *et al* 1966:35); it should not be cited separately, and it definitely does not contain an element ***-txo-**. This is another outrageous instance of arbitrary segmentation.

[101.] Bq **euli** 'fly' (n.)

Bq e-**uli** 'fly : Cauc (Archi) **hili**-ku 'fly ': Bur **ho:lal**-as 'butterfly, moth' (B93)

[102.] Bq **euri** 'rain' (n.): See under **ur** 'water'.

[103.] Bq **ez** 'not, no

Bq **ze**, **ez** (negative) : PNC ***ć**ə ~ ***ć'**ə (negative) : Bur **ačho** 'not yet' (B9ld) (B94b [13])

The Basque negative particle is universally *ez,* except that the earliest texts in B show a variant **ze** before subjunctive and imperative verb forms, a usage which survived at least until the late nineteenth century in places.

[104.] Bq **ezti** 'honey'

Bq **ezti** 'honey, sweet' : PNC ***mǐʒʒV** 'sweet', *h^{w}ǐ-**miʒʒū** 'honey' : Bur **mAč^{h}i** 'honey' (B9la [117])

Bengtson attributes all these to a PDC root ***miʒ* ~ ***ʒim** 'sweet'; observing the rather feeble resemblance to this of the Basque word, he proposes that this is derived from a prefixed and suffixed form along the lines of **hi-miʒ-ti*. This looks to me like a great deal of special pleading in support of an unconvincing etymology.

[105.] Bq **galdu** 'lose'

Bq **gal-** 'lose' : PNC *ʔi-**g^{w}Vl-** 'lose' (B9lc [87])

[106.] Bq **galtzar** 'side of the body'

Bq **gal**-tzar(-be) 'side, armpit' : PNC ***ʔā-G^{w}āɬ^{w}V** 'side ' : PY ***hol-** 'side (of body), cheek' (B9la [41]) (B9lc [109]) (R94b)

Basque *galtzar* means 'side of the body' and also in places 'arm' or 'chest'; only its derivative *galtzarbe ~ galtzarpe* means 'armpit' (*-be ~ -pe* 'below'). This word is a derivative of *galtza(k)* 'trousers', 'shorts', a conspicuous loan from Spanish *calzas* 'trousers, shorts'; see the Spanish word in Corominas and Pascual (1980).

[107.] Bq **ganga** 'palate'

Bq **gan**-ga 'palate', **gan**-gar 'uvula' : PNEC ***q̇wanʔu** 'face, cheek, flat surface' : PY ***KVn** 'face, mouth' (B9la [17])

Basque *ganga* means primarily 'arch, vault'; 'palate' is more usually the compound *ahoganga* (*aho* 'mouth'). Moreover, *ganga* is transparently a loan from Romance *canga*, whose central meaning is 'a yoking of two animals' but which has various transferred meanings. The derivative *gangar* could conceivably have been formed within Basque, but it is far more likely an independent loan from the Romance derivative *cangar*, attested in the sense of 'branch with a bunch of grapes' (compare Latin UVULA 'little bunch of grapes'). On all this, see Corominas and Pascual (1980) under *canga*.

[108.] Bq **gari** 'wheat'

Bq **gari** 'wheat', **garagar** 'barley' : PND ***k̆k̆ərV** / ***rək̆k̆V** 'wheat', 'rye', Proto-Lezgian ***karkar** 'wheat' (C85 [74])

A difficulty here is that Basque *gari* consistently shows a combining form *gal-*, as in *galburu* 'ear of wheat' (*buru* 'head'), *galbera* 'best wheat' (*bera* 'soft'), *galgorri* 'wheat (sp.)' (*gorri* 'red'), *galtzuri* 'wheat (sp.)' (*zuri* 'white') and *galbae* 'sieve' (*bahe* 'sieve'), among dozens of others. This suggests that *gari* is derived, quite regularly, from an original **gali*, which is not so convenient for C's purposes; see the remarks under **ukarai** 'wrist' below. The curious-looking word *garagar* has often been thought to be a reduplicated form of *gari*; if it is, it can add little to the proposed comparisons.

[109.] Bq **garkotx(e)** 'nape'

Bq gar-**khotx(e)** 'nape' (*gara* = 'skull') : Sum **kuš** 'side' (?) : PDC ****kuč** 'bone' : PW ****KATI** 'bone' (B9la [57 note]) (BR94 [7])

The word for 'nape' is undoubtedly a compound of some sort, but the authors' analysis of it cannot be accepted at face value. For one thing, no such word as **gara* 'skull' is attested; this is merely a fanciful notion of the lexicographer Azkue, and *gara* means nothing more than 'height', 'high part'. More importantly, the alleged **khotx(e)* 'bone' is unsubstantiated: no such word or element is found anywhere, and once again I must point out that ancient Basque words do not begin with voiceless plosives. It is far more likely that *-khotx(e)* here merely represents the Romance word for 'nape' (Occitan *cogòt*, Old Spanish *cocote*, etc.), which is also borrowed into Basque as *kokot(e)* ~ *kokots*: a formation like **gar-kokots* would have undergone the almost inevitable Basque haplology to yield **garkots*, whence the palatalized form *garkotx*. Further, the cited form for 'nape' is less widespread than the alternative *garkola* ~ *garkhora*, which shows the same first element but seemingly a different second element — most likely, Spanish *cola* 'tail'. Finally, the semantics suggested here are very odd: why should the nape of the neck, which is anything but a bone, be denoted by a compound meaning 'skull-bone'?

[110.] Bq **gau** 'night'

Bq **gau**, **gaba** 'night' : Proto-Lezgian ***xIam** (C85 .287])
Bq **gau**, **gab-** 'night' : Abkh **-Xa**, Proto-Lezgian ***X̄am** 'night' : PY ***saGar$_1$** 'spend the night' (B94a [17]) (B94b [28]) (SR94 [185])

The Basque word for 'night' is *gau*; the cited *gaba* is merely a western variant of its definite form *gaua*, and strictly secondary.

Bouda (1948:§160) compares Basque *gau* with both North and South Caucasian

[111.] Bq **gela** 'room, chamber'

Bq **gela** 'room', 'habitation' : PND ***qəlV** 'house, dwelling' (C85 [68])

The Basque word *gela* means 'room', 'chamber', but not 'habitation', and it is a transparent loan word from Latin CELLA 'chamber' (Michelena 1977:320), a possibility allowed for by C himself in his entry.

[112.] Bq **gibel** 'liver'

Bq **gibel** 'liver' : PNC ***ƛ'ʷäHälV** 'liver' (B9lb [9]) (B9lc [71]) (B9ld PS1)

[113.] Bq **giltza** 'key'

Bq **kilte** 'key' : PND ***k'ule** (C85 [66])

There is no such Basque word as **kilte*; the universal Basque word for 'key' is *giltza*. The correct form is still no doubt adequate for C's purposes, but this constant citing of erroneous and non-existent forms can do little to inspire the reader's confidence. There is also the word for 'nail', which is *itze* in the east but *iltze* or *ultze* in the west; Michelena (1977:74) considers that *iltze* represents a contamination of original *itze* by the unrelated *giltza*.

Bouda (1949:§138) also compares Basque *giltza* and *itze* (which he considers to be related) with North Caucasian forms.

[114.] Bq **giltzurrin** 'kidney'

Bq **gul**-tzurrin ~ **gil**-tzurrin 'kidney' : PNC ***q'ʷɔ̄l-(V)q'ʷV** 'egg, seed, grain' (B9la [49])

The most widespread of the several variants of the Basque word is *giltzurrin*. The first element here is surely *giltz(a)*, which means 'joint' in isolation, but which in compounds sometimes appears to mean more specifically 'hip' or 'pelvis' (*giltzagain* 'ilium', from *gain* 'top'). The second element is obscure: neither *urrin* ~ *urrun* 'far' nor *urdin* 'blue' (but formerly 'green') seems obviously right. In any case, a fanciful derivation from a hypothetical word for 'egg' is too far-fetched to be taken seriously.

[115.] Bq **gizen** 'fat' (adj.)

Bq **gizen** 'fat' : PY ***giʔd** 'fat' (n.) (SR94 [88])

[116.] Bq **gizon** 'man'

Bq **giz** 'man, person' : Abkh **-ҳać'a** 'man, person' (C85 [63])

A curious error: the Basque word for 'man' is *gizon*, not **giz*, which does not exist. The Basque word is recorded as *CISON* ~ *CISSON* in Aquitanian texts (Michelena 1964a:18). The rather feeble Abkhaz match does not even appear to have any Caucasian standing.

[117.] Bq **gogor** 'hard'

Bq **gogor** 'hard' : Abkh **RʷRʷa** 'strong', 'firm' (< PAA ***-GʷəGʷa**), PND ***q'q'waInq'wV** 'hard' (C85 [43])

Bq **gogor**(r) 'hard, stiff, firm, cruel' : PNC ****ğğʷV̆rV** 'hard, severe, dangerous' : Sum **gur** ~ **gür** 'great, mighty' (B9la [114])

Bq **gogor** 'hard' : PNC ***GʷērV** 'stone' : Bur **γoro** 'stone' (B9lc [108])

Michelena (1970b:72) sees Basque *gogor* as deriving from an earlier **gorgor*, a reduplication of *gor*. This last is today the usual word for 'deaf', but it is also attested as meaning 'insensible' and 'unyielding', providing some

semantic underpinning to Michelena's proposal. For once, then, it appears that the Basque form is too short to match the Caucasian words, or some of them — the authors appear to be very uncertain which Caucasian words they want to compare with the Basque.

[118.] Bq **golko** 'space between one's chest and one's clothes', 'bay, gulf'

Bq **kolko** 'breast (of a woman)' : Abkh **-k'ək'a-** 'breast (of a woman)', PND ***kk̄/x̄kk/u/** 'nipple', 'breast' (C85 [44])

Bq **kholko** ~ **golko** 'breast, bosom' (from ***khoklo** ?) : Abkh **-k'ək'a-** 'female breast', ECauc ***kx̱k[u]** 'breast, nipple' (B94a [24])

Basque *golk(h)o* ~ *k(h)olko* does not mean 'mammary gland', which is variously *ugatz*, *bular*, or *diti*. It means 'breast', 'bosom' (of a man or a woman), but not in the anatomical sense: it rather denotes the space between one's chest and one's clothing, or sometimes, in the case of a woman, the space between her breasts, and it is most commonly used with reference to carrying or hiding something in this location. It also has another meaning: 'gulf', 'bay' (a rounded inlet of the sea), as in *Bizkaiko Golkoa* 'the Bay of Biscay'. The presence of the voiceless plosive after *l* (native words have only voiced plosives in this position) shows that it is a loan word, and in fact it is a borrowing from Latin COLPU, itself borrowed from Greek κόλπος (Michelena 1974:193) and is therefore of the same origin as English *gulf*, borrowed from Old French.

[119.] Bq ****gor; gorri** 'red'

Bq **gor** 'meat', **gorri** 'red', **gor**-din 'raw, crude', **gorrin**-ko 'yolk of an egg : Sum **gur** ~ **gurun** 'blood' (B9lc PS2 [37])

There is no such Basque word as the alleged **gor* 'meat'; this is merely another of Azkue's little flights of fancy, like **gara* 'skull' above. The adjective *gordin* is securely derived by Michelena (1970b) from *gor* 'deaf', but formerly 'firm, unyielding', by means of the familiar adjective-forming suffix *-din*. There is no earthly reason to relate *gorri* to these other words, though *gorringo* (to cite its most usual form) is an obvious derivative of *gorri*.

[120.] Bq **gose** 'hungry'

Bq **gose** 'hungry, hunger', **gose**-te 'famine' : PNC ***g(g)aši** 'hunger' (B9la [119]) (B9lc [66], [88]) (R94b)

The word *gosete* is a transparent and regular derivative of *gose* and should not be cited independently.

[121.] Bq **gozo** 'sweet'

Bq **gozo** ~ **goxo** 'sweet, tasty' : WCauc ***q'aş'ʷə** 'sweet', ECauc ***q'(w)Vlć'ć'V** 'sour' : Bur **gaš**-ar-um 'salt-sweet' : Sum **kuš$_6$** 'honey, sweet' (B94a [30])

Basque *goxo* is merely an affective form of *gozo*.

[122.] Bq **gu** 'we'

Bq **gu** 'we' : PND ***-γü** 'we (inclusive)' (C85 [12])

Bq **gu** 'we' : Cauc ***ǯ(ü)** 'we' (inclusive) (B94b [2])

I presume Čirikba and Bengtson are citing the same Caucasian form here. Many others have preferred to point to a connection with the Kartvelian first-plural agreement marker *gw-* ~ *gv-* (Uhlenbeck 1924:578; Trombetti 1925:86; Lafon 1944:529; Lafon 1952a:§24). Lafon also points out the match with Northeast Caucasian.

[123.] Bq **(g)une** 'stretch, interval', 'moment, occasion'

Bq **une** 'place' : Abkh **-nə** 'place, country', Bzyb Abkh Apś-**nə** 'Abkhazia' (cf. also Bq, Abkh **-n** locative suffix) (C85 [80])
Bq (Z) **gune** 'place' : PNC ***GʷīnhV** 'village, hut' (B91c [111], PS1)

It is a gross oversimplification to assert that Basque has a word *une* or *gune* meaning 'place'. In fact, this item exhibits decidedly complex behavior. The easternmost varieties show a word *gune* (the Z form is in fact *güne*) meaning 'space, room; stretch, extent; interval'. Some dictionaries list 'place' as one of the possible translations, but this is somewhat misleading: the adverbial derivative *guneka* can indeed be rendered 'in places', but a more literal rendering would be something along the lines of 'at intervals, here and there, now and again'. In most other varieties, the word appears in the form *une*, and it means either 'moment' or 'occasion'. In a small area in the north, the word again appears as *gune* but with the surprising meaning of 'gesture', 'expression'. Finally, the same morpheme is widely attested as a derivational suffix *-(g)une*, and it is only in this use that a gloss of 'place' seems at all appropriate: *itzalgune* 'shady place' (*itzal* 'shade'), *ibilgune* 'road' (*ibili* 'travel'), *auzune* 'neighborhood' *(auzo* 'neighbor'). But all such derivatives carry a clear sense of 'extension in space or time'. Those with a notable temporal sense include *ezune* 'time of scarcity' (*ez* 'no, not'), *isilune* 'period of silence' (*isil* 'silent'), and *oraigune* 'a short time ago' (*orai(n)* 'now').

It seems, then, that the likeliest original meaning for this item was something like 'extension in time or space' or 'interval', and not 'place'. This is not helpful to Čirikba or to Bengtson. As so often, these authors do not agree as to which Caucasian "cognates" they identify, and neither 'place, country' nor 'village, hut' looks like a very persuasive match. Furthermore, we must surely take **gune* as the original form of the Basque word, and not **une*, since there are many parallels in Basque for the loss of initial *g-*, but hardly any for the insertion of an initial *g-*. Finally, it should be noted that *(g)une* is also phonologically anomalous, with its intervocalic *n*. The obvious way out is to posit an original **gunne* (Michelena's **guNe*), which would yield the required *gune*, but Michelena (1977:305) points out that, in place names, this element regularly appears as *-(g)ue*, showing the expected loss of intervocalic *n*, and leaving this item very mysterious indeed.

[124.] Bq **gurin** 'butter'

Bq **gorhi** ~ **guri** 'butter' : Sum **gár** 'lait crémeux' (B9lc PS2 [8])

The earliest attested sense of Basque *gurin* and its variants is 'lard, solid animal fat'; today, it commonly means 'butter', but in places also 'clotted cream' or 'juice' (of vegetables or meat). The matchup seems implausible and accidental.

[125.] Bq **gurpil** 'wheel', **inguru** 'vicinity', 'around, near'

Bq **gurpil** 'wheel', **ingurru** 'around' : Abaza **gʷər gʷər**, Abkh **-gʷərgʷəl** 'ring', etc. (C85 [67])

Čirikba is apparently under the impression that these two Basque words are related. They are not. First, *gurpil,* originally 'cartwheel', but today 'wheel' in general, is a transparent compound of *gurdi* 'cart' and the element *-bil*, which appears as the final member of a number of compounds; the development of **gurdi-bil* into *gurpil* is absolutely regular. The element *-bil* occurs in several formations and generally carries the notion of roundness: *barrabil* 'testicle', *ukabil* 'fist', *luharbil* 'clod of earth', *boro(n)bil* 'globe, sphere', *biribil* 'round'; this last appears

to be a reduplicated form **bilibil* (Michelena 1973:§162). The element *gur-* in *gurpil* thus has nothing to do with roundness.

The word *inguru* (not **ingurru*) means 'vicinity'; its locative form *inguruan* is a common postposition meaning 'around, in the vicinity of'. The word is borrowed from Latin IN GYRUM or IN GYRO 'in a circle', of Greek origin (Michelena 1977:79-80).

[126.] Bq **hagin** 'molar'

Bq **hagin** 'molar' : PW ***GINI** ~ ***NIGI** 'tooth' (BK93)

[127.] Bq **haitz** 'rock, crag'

Bq (h)a-**itz** 'rock' : PY ***čɨʔ**-s 'stone' : Sum **zá** ~ **zé** 'stone' (B9la [80])

There is some evidence for a lost nasal in *haitz*, and Michelena (1949b:211) therefore reconstructs **anitz* or **anetz*. Though apparently unaware of this, Bengtson, in a footnote to his comparison, himself admits that the Basque word is probably not related to the Yeniseian and Sumerian items after all.

[128.] Bq **hamar** 'ten'

Bq **hama**-r 'ten' : PNC ***Xa̲mɦV** 'handful' : Sum **haw(a)mu** ~ **haw(u)mu** 'ten' (B9lb [6]) (B9lc [115], PS2 [28])

The *m* in *hamar* cannot be original. We might expect an original **anbar*, but there is no direct evidence for this. Tovar (1958) attempts to relate this word to *amai* 'end', but Michelena (1972:82) rejects this, seeing *amai* itself as derived ultimately from a compound of *hamar*.

[129.] Bq **handi** 'big'

Bq (h)an-**di** 'big' : Abkh a-**du** 'big', Ubykh jă-**dă** 'many' (C85 [38])

Here Čirikba is apparently segmenting the entire first syllable into oblivion in order to extract the single segment for which he wants to claim a match. This is just the sort of thing for which Michelena scolded Bouda, and I take an equally dim view of it. Michelena (1954) suggests that this adjective is present in a range of Aquitanian personal names, such as *Andose* and *Andossus*.

[130.] Bq **haragi** 'meat, flesh'

Bq **aragi** 'flesh' : PAA ***ɣ̌a**, PND ***rük̆'k̆'V** 'flesh' (C85 [76])
Bq a-**ragi** 'flesh' : PNC ***ra̲ă̲'ɨ̲** 'flesh' (B9la [44]) (B9lc [72])

There is considerable internal evidence in Basque that *haragi* is a derivative of an earlier **ara* 'meat, flesh' : Z *aratsu* 'fleshy' (*-tsu* 'full of'), HN LN *aratxe* 'veal', and a few others. If this interpretation is correct, the matchup is destroyed. Note also that Bengtson, who nearly always treats Basque *h* as etymological and matches it with segments in the other languages, in this case does not do so and does not even note the existence of the aspirated form *haragi,* even though this form is virtually universal in the aspirating dialects.

Bouda (1948:§27), who considers *-gi* to be a suffix in this Basque word, cites some rather different parallels with North Caucasian languages.

[131.] Bq **harri** 'stone, rock'

Bq **harri** 'stone' : Bur **kar**-kat, γa-**k(h)arr** 'stony ground' (B9lb [17]) (B9lc [84])

Interestingly, Basque *harri* is one of the very few cases in which there is reason to believe that the initial *h* might represent an original **k*, since an element *karr-* meaning 'stone' is widely attested in Romance and appears to be of substrate origin; see Michelena (1977:219).

[132.] Bq hartz 'bear'

Bq **hartz** 'bear' : (Cauc) Dag ***XIʷVrćV** 'squirrel, marten', Dargwa **X̄aIrc'** 'squirrel' : Agul **Hürč**-al 'marten, hamster', Nakh ***Xešt**, Chechen, Ingush **Xešt** 'otter' : PY ***Xas** 'badger' : Sum **az** 'bear' (B9la [90]) (B9ld PS1) (B9le)

The semantic parallels offered by Bengtson are rather unexpected, to say the least — 'bear' and 'hamster'? — though he describes them as "plausible in light of time depth and population movements". In any case, Basque *hartz* has long been thought to be a loan word from Indo-European — possibly from Celtic, though the presumed Celtic **artos* is not phonologically ideal, more likely from a more conservative form like **arktos* or **arksos*. See Agud and Tovar (1988-) or Gorrochategui (1987) for discussion, and see under **azkoin** 'badger' above.

[133.] Bq **hats** 'breath'

Bq **hats** 'breath' : (Cauc) Chechen **ħožu** 'odor', Batsbi **ħaič'** 'odor' : Bur **hiš** 'breath' (B93)

[134.] Bq **hatz** 'finger', 'paw (of an animal)', **behatz** 'toe'

Bq **atz, beatz** 'finger' : Abkh, Ubykh **-cʷa** 'finger' (C85 [50])

Bq **hatz** 'finger, claw, paw', be-**hatz** 'toe, thumb, claw, hoof', etc. : PNC ***kʷač'e** 'paw', ***kʷašV** 'paw', ***k'VsV** 'finger' : Bur **qAṣ** 'length from elbow to fingertips, cubit, half-yard' : PY ***ki?s** 'foot' (B9la [31l) (B9lc [52], [81], PS1) (B9ld PS1) (B93) (B94b [15]) (R94b)

Basque *hatz* does not mean 'claw'; it means 'finger' (of a person) or 'paw' or 'leg' (of an animal). Many varieties have *behatz* for 'finger', instead of *hatz*. The explanation, however, is clear, and it has nothing to do with an alleged "body-part prefix". Western varieties make a distinction between *hatz* 'finger' and *behatz* 'toe', showing that the element *be-* in the second word is merely the common item *behe* ~ *be* 'lower, below'. Other varieties have often lost the distinction between these two words, possibly under the influence of Spanish, in which the word *dedo* is used indifferently for 'finger' and 'toe'.

Bouda (1949:§59) points out precisely the same resemblance that C adduces, and adds some more remote similarities from other Caucasian languages.

[135.] Bq **hau** 'this', **han** '(over) there', **hara** 'thither'

Bq **au** 'this' : Abaza **a-u**-j 'that', Abkh **u-j** 'this', **wa** 'there', Adyghe **au(a)** 'there', PND ***u̯ĕ** 'this', 'that', etc. (C85 [14])

Bq **an** 'there', **aña** 'as much as' : Abkhaz-Tapantan **ana** 'there', PND ***nV** 'this', 'that', etc. (C85 [15])

Bq **ara** 'thither', **ara**-ño 'up to there', **ara** 'there (it) is', or 'there', **ori** 'this', **ura** 'that', etc. : Abkh **ar**-i 'this', **ara**-ś 'hither', **ur**-t 'those, they', etc. (C85 [16])

There are several Basque stems tangled up in these last three entries. Basque has three demonstratives, each of which shows stem suppletion: proximal *hau, hon-* 'this' (a few varieties have *haur,* which is certainly the older form), mesial *hori, horr-* 'that' (not 'this'), and distal *hura, har-, hai-* 'that (over there)'. The forms *han* 'there', *hara* 'thither', and *haraino* merely represent the stem *har-* of the third demonstrative with the ordinary case suffixes

-n (locative, discussed below), *-a* (allative, discussed below under **-ra**), and *-aino ~ -año* 'as far as, up to'. The form meaning 'as much as' is more conservatively cited as *adina*; this form is not straightforwardly related to the demonstrative stems (see Michelena [1970b] for a study of this and related forms). The occurrences of *n* in the Basque forms in C's item [15] are all case endings and not part of the stem; hence the Proto-Nakh-Daghestan form is irrelevant. In any case, most of these resemblances are noted by Dumézil (1933:133-135) and by Lafon (1951:§§11-14).

[136.] Bq **hau(r)** 'this'

Bq **-haur** 'self' : Bur **-khAr** 'self' (B91c [82])

Basque *-haur* does not mean 'self'. Bengtson has extracted this morph from the intensive pronouns like *nihaur* 'I myself', under the mistaken impression that *ni- is* a possessive element (199la:75), but the first element is simply the pronoun *ni* 'I' while the second is merely the proximal demonstrative *hau(r)* 'this'; all the Basque intensive pronouns are formed in this manner.

[137.] Bq **hauts** 'dust, powder'

Bq **hauts** 'dust' : PNC ***XurtV** 'foam, scum': Bur **xUrc ~ xuurc** 'dust' (B91a [77]) (B91c [112])

[138.] Bq **hauzo** 'neighbor', 'neighborhood'

Bq **hauzo** 'neighbor' : PNC ***hV̲č'wV** 'guest' : Bur **ašo** 'guest' (B93)

[139.] Bq **hazi** 'grow, grow up'; 'raise, bring up; cultivate'

Bq **azi** 'grow, grow up' : Abkh **a-z-ha-ra** 'grow', PND ***VǯV-** (C85 [32])

The Basque verb *hazi* appears to be derived, via the ancient verb-forming suffix *-i*, from the noun *hatz* 'lineage, race' (Michelena 1977:289), an interpretation which is not helpful to Čirikba. The Basque-Abkhaz match is cited by Bouda (1948:§38).

[140.] Bq **hegal** 'wing'

Bq he-**gal** 'wing, fin' : Bur **gAI**-gi 'wing, fin' (B9la [32]) (B91c [89])

This is another case in which Bengtson drops the initial *h-* of Basque without comment.

[141.] Bq **hegi** 'ridge'

Bq **hegi** 'peak' : PNC ***ʕerqwē** 'mountain' : Bur **hu:rgo** 'slope, uphill' (B93)

Basque *hegi* does not mean 'peak', but rather 'ridge'; in Z, it also means both 'edge' and 'corner'. A sense of 'hill' appears to be attested in place names (Michelena 1977:125), though not for the independent word.

[142.] Bq **herri** 'country, town, inhabited place, people'

Bq **erri** 'people' : PAA ***ɬə** 'army', PND ***ʔIwə̌rʔV** 'people', 'army' (C85 [62])

The Basque word *herri* means both 'inhabited place' and 'people who live in a particular place', regarded collectively, most often the speaker's own compatriots. It seems impossible to judge which of these related meanings is the earlier: Azkue favors the first (as do I), while Michelena favors the second. The word does not mean anything like 'crowd, host, multitude', let alone 'army' .

Bouda (1948:§122) prefers to compare this word to words in South Caucasian languages.

[143.] Bq **hertze** 'intestine'

Bq **(h)ertze** ~ **(h)erxe** 'intestine' : PNC ***Hĕ-wĕrk̆'k̆'V** ~ ***-p̄ĕrk̆'k̆'V** ~ ***-bĕrk̆'k̆'V** 'large intestine' : Bur **wArk** 'fat on the big intestine of animals; a kind of sausage' : PY ***pi?il'** 'intestine' : Sum **peš** ~ li-**beš** 'heart bowel, womb' (B9la [54])

[144.] Bq **heze** 'moist, damp'

Bq **eze** 'damp' : Abkh **aʒa** 'damp' (cf. **a-ʒə** 'water, river') (C85 [34])

Basque *heze* in fact means 'moist' only in the eastern dialects; elsewhere it means 'flourishing, succulent' (as applied to plants) as opposed to 'dried-up, withered'. Exactly the same parallel is cited by Bouda (1948:§37).

[145.] Bq **hi** 'you' (singular intimate)

Bq **hi** ~ **hi-** ~ **he-** 'thou, thee, thine'; **h-**, **-k** (second person verbal elements) : PNC ***Gu** 'thou' : Bur **gu-**, **go-** 'thou, thy', **guu**-i 'thou thyself', **goo**-r 'to thee, for thee' : PY ***kV-** / ***?Vk-** 'thou' (B9la [143]) (B9ld) (B94b [4])

As usual, there is no reason to suppose that Basque second-person *h-* is etymological. On the other hand, it is certain that the agreement suffix *-k* (male) derives from earlier **-ga*. Now I want to draw attention to this entry, because it well illustrates a key point. The second-person singular agreement markers in Basque exhibit interesting and seemingly ancient alternations: prefix *h-*, suffix **-ga* (male) but **-na* (female) (> *-n)* (it is possible that **-na* derives from **-naga*; see Jacobson 1975). If we were really looking at traces of an ancient genetic unity, then we might expect evidence of similar alternations in some of the other languages adduced — especially since this sex distinction is completely isolated and anomalous in Basque. But Bengtson adduces no such evidence, and I can only conclude that none exists. Like all the earlier work dismissed by Michelena, therefore, the "Macro-Caucasian" hypothesis fails precisely at the point at which a valid hypothesis might have been expected most particularly to provide illumination.

[146.] Bq **hil** 'dead', 'die', 'kill'

Bq **il** 'die, kill' : PAA ***k̆''V** 'kill', PND ***?i(w)k̆'V** 'die' (C85 [29])

Bq **hil** 'dead; die', **herio** 'death' : WCauc ***x̄'V**, ECauc ***?i(w)x̄'V** 'die, kill' : Bur **hōl** 'army' : Sum **hul** 'destroy' (B91c PS2 [16]) (B94a [14]) (B94b [20])

Čirikba and Bengtson are in fact citing the same Caucasian forms, in spite of their differing transcriptions. It is clear that Basque *hil* is an adjective in origin, and that its verbal use is secondary. It is far from clear whether *herio* is connected with *hil*; a connection with the verb *ero* 'kill' looks slightly more plausible, but would still be morphologically opaque.

[147.] Bq ***hile-** 'moon'

Bq **(h)il** ~ **(h)ila** 'moon', **hil**-(argi), (LN) **il**-azki 'moon, moonlight', **hila**-(bethe) 'full moon, month' : Bur **hAI**-Anc ~ **hAIA**-ns ~ **hAI**-Anz 'moon' (B9la [72]) (B91c [79]) (B91d) (B94b [22])

It is out of order for Bengtson to cite an alleged Basque **(h)il(a)* 'moon', with no asterisk and without even a hyphen, since no such word is attested anywhere. The existence of *(h)ilargi* 'moon' (*argi* 'light') and of *(h)ilebet(h)e ~ (h)ilabete* 'month' (*bete* 'full') leads Michelena (1977:411) to reconstruct **ile* or **iLe* for 'moon'. The Low Navarrese form, cited from Azkue (1905) as "*ilazki*", is an error: the correct form is *ilaski,* with a mysterious second element.

[148.] Bq **hilindi** 'charcoal'; 'something scorched'; 'blight'; 'dead stalk of wheat'; 'rye grass'; 'firebrand'; 'spark'

Bq i-**lhinti** 'firebrand, ember' : PNC ***ɬwind** 'firewood' (B93)

Basque *hilindi ~ ilhindi ~ ilhinti ~ ~ ilhintx ~ itxindi* has an extraordinary range of meanings variously centering around the notions 'dead', 'black', 'burnt' or 'burning'. It is hard to know what to make of this, but the first element does bear a striking resemblance to *hil* 'dead'. There is no particular reason to assume that the sense singled out by Bengtson is the earliest meaning of the word. In any case, the presence of the variant *ilinti* in the west, with its *-nt-* cluster, shows indisputably that this word cannot be ancient.

[149.] Bq hitz 'word'

Bq **itz** 'word' : Abkh **cʷa**, Avar **-ic-** 'speak', etc. (C85 [60])

Bouda (1949:§40) draws attention to resemblances between the Basque word and words in several Caucasian languages, including Avar.

[150.] Bq **hogei** 'twenty'

Bq **hogoi** 'twenty' : PNC ***Gə** 'twenty' : PY ***χɔGa** 'ten' (SR94 [270])

Northern *hogoi* is merely an assimilated variant of the more usual *hogei.*

[151.] Bq **horma** 'ice'; 'wall'

Bq **horma** 'ice' : Bur **γAmu ~ gAmu** 'ice' (B9lb [22])

The Basque word, which means both 'ice' and 'wall', is a loan from Latin FORMA (Michelena 1977:51). It is attested in early texts in the expected form *borma*, but it has undergone the virtually regular loss of initial *b-* before *o*, followed by the regular insertion of unetymological *h-*.

[152.] Bq **hortz** 'incisor, tooth'

Bq **hortz** 'tooth' (by metathesis) : PNC ***cɨɬ̃V** 'tooth', ***gwəlǯwē** 'fang, canine tooth' : Bur **-hAše-**(me) 'molar' (B9lc [122]) (B9ld) (B94a [11]) (B94b [12])

[153.] Bq **hotz** 'cold', **izotz** 'frost, ice'

Bq **otz** 'cold', **izotz** 'frost' : PAT ***c'a-**ʕa 'cold', PND ***Vc'c'Vr-** 'freeze' (C85 [82])
Bq **hotz** 'cold' : (Cauc) Avar **kʷač** 'coolness; frost' : (Yen) Pumpokol **ki:č**-idin 'cold' (B94b [34])

Basque *izotz* 'frost' is a compound whose second member is *hotz* 'cold' and whose first member is probably *ihintz* 'dew' (Michelena 1977:411). As so often, Čirikba and Bengtson differ in their choice of Caucasian cognates, with

the former treating Basque *h* (correctly) as non-etymological. The same observations are made by Bouda (1948:§50).

[154.] Bq **hur** 'hazelnut, nut'

Bq **ur** 'nut' : Abkh **a-ra** 'nut tree', **a-ra**-ssa 'small nut' (C85 [73])
Bq **hurr** ~ **ur** 'nut' : PNC ***hʷɔ̄r-ƛ'V** 'nut' (B9la [105]) (B9lc [120])

The Basque word *hur* means more precisely 'hazelnut', though it is sometimes used as a generic label for 'nut' in general. The same comparison is drawn by Bouda (1948:§107).

[155.] Bq **-i** (dative case-suffix)

Bq **-i** ~ **-ri** (dative) : PNC ***-i** / ***-j** (dative) : Bur **-e** (genitive, general oblique) (B9ld) (B93)

Bengtson suggests that the morph *-ri* results from the compounding of dative *-i* with another case-ending, but this is not so. The *-r* comes about as follows. The distal demonstrative stem is *(h)ar-*: hence, for example, *gizon hari* 'to that man' (dative). As explained above under **-a**, this demonstrative has developed into a phonologically bound article, and hence we have forms like *gizonari* 'to the man'. But, in the absolutive **gizon har* 'that man', the final *-r* was lost, yielding *gizona* 'the man'. The resulting pattern *gizona / gizonari* led to reinterpretation of the *-r-* as a morph preceding certain case suffixes, and as a result this *-r-* has been analogically extended to forms in which it was not historically present.

[156.] Bq **ibar** 'water-meadow', 'valley'

Bq **i-bar**(r) 'valley' : Bur **bAr** 'nullah, ravine, valley' : Sum **bar**, gu-**bár**-ra 'open field, wilderness' (B91a [81]) (B9lc PS1)

Today Basque *ibar* most commonly means 'valley', a sense in which it competes with the other word for 'valley', *haran*. However, there is abundant evidence that *ibar* earlier meant the same as Spanish *vega*: 'water meadow', 'fertile low-lying land along a river'. (In the Basque Country, with its narrow V-shaped valleys, such terrain is found nowhere but in valleys.) There is no doubt that this word is related to *ibai* 'river'; it may further be related to *ibi* 'ford'; 'low water, low tide', though this seems very unlikely. There have been strenuous efforts to relate *ibar* also to *barren* 'interior part', 'lower part', but these efforts cannot account for the initial *i-* (which Bengtson, as always, dismisses as an inconsequential "prefix" of some sort).

[157.] Bq **idi** 'ox'

Bq **idi** 'ox' : Abkh **-tə** 'ram', Abaza **tə** 'castrated ram', Ubykh **t'ə** 'ram', etc. (C85 [69])

Since 'ox' and 'ram' hardly constitute a persuasive semantic match, Čirikba's intention is presumably to link these words via the sense of 'castrated animal'. Trombetti (1925:§181) proposes to link Basque *idi* to a Chechen word which actually means 'ox', though Bouda (1949:55) is critical of this on phonological grounds.

[158.] Bq **igel** 'frog'

Bq **i-gel** ~ **in-gel** ~ **i-hel** ~ **ne-gel** ~ **ne-gal** 'frog' : PNC ***q'q'ʷVrV**-q'q'V 'frog' : Bur **γUr-kUn** ~ **gUr-qUc** 'frog': PY ***xiʔr-** 'frog' (B9la [95]) (R94b) (SR94 [106])

By far the most widespread (and earliest attested) form of the Basque word is *igel*. It has often been thought by vasconists that this might be a derivative of *igeri* 'swimming, floating', but this derivation is not certain.

[159.] Bq **ihintz** 'dew'

Bq i-**hi(n)**-tz 'dew', **u-hin** 'wave' : PNC ***xə̲nʔi** 'water' : Bur **huu-š** 'wetness (of ground), moisture' : PY ***xur**$_1$ 'water' (B9la [84]) (B9lc [96]) (SR94 [290])

The comparative evidence, with nasalized vowels in the east, points unmistakably to an original **initz* or **inintz* for 'dew', which destroys the match (Michelena 1950a:449; 1977:305). For the unrelated Basque *uhin*, see **uhain** 'wave' below.

[160.] Bq **-ik** (partitive suffix)

Bq **-ik** (abstract or indefinite, suffix) : Abkhaz-Tapantan **-k'** (indefinite and singulative), Ubykh **-k'ə** (singulative), Abkhaz-Tapantan **ak'a** 'one', etc. (C85 [18])

The Basque suffix *-ik* is not well described as "abstract" or "indefinite": it is a *partitive* affix occurring chiefly in negative and interrogative sentences and functioning much like English 'any'. In certain circumstances it appears to mean something along the lines of 'from', 'out of', and it quite possibly derives from an ancient ablative, but it never has a meaning resembling 'one', and the Caucasian parallels therefore appear to be insubstantial. Recognizing the semantic difficulty, C attempts to rescue matters by identifying partitive *-ik* with his putative **ika* in the next entry, to which I now turn.

[161.] Bq ****ika** 'one'

Bq ***ika** 'one' (in *amar-ika* 'eleven', i.e. '10 + 1'): Abkh **ak'ə**, Abaza za-**k'ə**, Ubykh **-k'ə** (< **k'ə* 'one') (C85 [19])

The numeral *hamaika* 'eleven' (never **(h)amarika,* as suggested by Čirikba) has, as Michelena (1977:496) points out, no etymology. The first element is undoubtedly *hamar* 'ten'. The second element, which Michelena (1977:117) concludes, after examining the dialectal variants, is probably **-eka*, rather than **-ika*, is completely unique. C's confident gloss of 'one' is no more than a wild guess: the universal Basque numeral for 'one' is *bat*, which, moreover, itself appears to be embedded in the numeral for 'nine' (see ****-tzi** below), suggesting that *bat* is very ancient and that **-eka* could not have meant 'one'. It is perhaps worth pointing out that *hamaika*, apart from meaning 'eleven', is also the word used throughout the Basque Country to denote an indefinitely large number, rather like English 'zillion'. Given the utterly transparent formation of the Basque numerals from 'twelve' upwards, I am inclined to wonder whether *hamaika* might not represent a fossilized relic of the time when the Basque counting system stopped at 'ten', with an original meaning along the lines of 'ten-something', i.e., 'lots'. Of course, this is unbridled speculation, but so is C's 'one'. A further difficulty is discussed under ****-tzi** below.

Salmons (1992) points out that the alleged Basque **ika* could be associated with PW ****TIK** 'finger; one', a matchup ignored by Bengtson and Ruhlen (1994 [23]) and not supported by Salmons either.

[162.] Bq **ile** 'hair'

Bq **illea** 'wool', **illeak** 'hair' : Abkh a-**lasa** 'sheep's wool', PAA ***laśʷV** 'wool' (C85 [75])
Bq **ilhe** ~ **ulhe** 'wool' : PNC ***ʔəlXIV** 'wool' (B9lc [116])

The primary meaning of Basque *ile* today is 'hair', though 'wool' is also well attested; the sense of 'wool' is more commonly expressed by the compound *ardi-ile* or *artile* (*ardi* 'sheep'). This word exhibits an unusual range of dialect variants: central dialects have *ile*, *ilhe*, *ille*, *eile*, or *elle*; Bizkaian in the west has *ule* or *ulle*; Zuberoan and

Roncalese in the east have *hule* or *ule*. While it is familiar for Bizkaian to have *u* where other dialects have *i*, it is very unusual for the eastern dialects to share this *u*. Michelena (1977:73-75) reconstructs **iLe* ~ **uLe*, and then reaches the conclusion that the original vowel in this word must have been *u*, and that this must have undergone fronting in the central dialects under the influence of the following palatal lateral, which appears to be ancient. He also delicately considers the possibility of a lost initial labial consonant, and cites without comment the proposal of Uhlenbeck (1940-41) that the Basque word was borrowed from an Indo-European language: note for example Gothic *wulla* 'wool' (Basque was in contact with Gothic during the period of the Visigothic kingdom in Spain).

[163.] Bq ****ilu**

Bq **ilu** 'move' : Abkh **-la**-ra 'go', etc. (C85 [27])

This is mysterious. No such word as **ilu* exists in Basque in this or any other meaning, and no verb meaning anything like 'move' has a form even approximating to C's *ilu*, the closest match being *ibili* 'move about, be active'. In fact, *ilu* is not even a possible form for a Basque verb; moreover, if such a bizarre form had ever existed, the intervocalic *l* would not have survived.

[164.] Bq **inguru** 'vicinity': See under **gurpil** 'wheel'

[165.] Bq **intzaur** 'walnut'

Bq **intzaur**(r) ~ **intxaur**(r) ~ **eltzaur**(r) 'nut' : PNC ***Xüwx̄(ʷ)o** 'nut' : Bur **xUnzUr** 'kernel of a walnut' : Sum **nunuz, nuz** 'egg' (B91a [106]) (B91c PS2 [46])

The Basque word means specifically 'walnut', and it is surely a compound of *hur* 'nut' with an unidentified first element. The parallels adduced are less than overwhelming.

[166.] Bq **intzigar** 'hard frost'

Bq in-**tzig**-ar 'frost' : PNC ***čʰowq̄ʰIV** 'sleet, hoarfrost' (Avar **tˢiq** 'sleet') : Bur **čʰAγ**-(uurUm) 'cold' (adj.) : PY ***caG** (Kot **sāk** 'icy crust on snow') or ***tiχ** 'snow' : Sum **še** 'cold, frost', **šèĝ** 'rain, (to) rain' (B91a [86]) (R94b)

Western Basque *intzi(g)ar* 'hard frost' is analyzed as containing an alleged root **-tzig-* 'frost'. But this word is merely a local variant of the more usual *aintzigar*, whose second element is transparently western *igar* 'dry' (< *eihar*, still today the eastern form), and whose first element is quite possibly *ihintz* ~ *intz* 'dew'. If it is not, we must turn to the widespread *aintzira* 'swamp, lake' and the localized *aintzika* 'reservoir', which appear to be variants or derivatives of *intzura* ~ *intzira* 'swampy place', a word derived by Michelena (1973:99) from Latin INSULA 'island'.

[167.] Bq **ipini** 'put'

Bq **min-** 'put, place, arrange, set up, settle' : PNC ***ʔi-man-** 'stay, be' : Bur **-mAn**-(As) 'be, become' : Sum **me** 'be' : Hurrian **mann-** 'be': Urartean **man-** 'be' : PW ****MANA** 'stay (in a place)' (B91b [62]) (B91c [13], PS2 [41]) (BR94 [14])

There is no such word as **min-*; what Bengtson is trying to cite is *imiñi*, a regional variant of the more widespread *ipini* ~ *ibeni* (and others) 'put' (it doesn't have the other meanings imputed to it by Bengtson and Ruhlen), in which the *m* results merely from the familiar Basque process of nasal assimilation, once again destroying the match.

[168.] Bq **irazi** 'strain, filter'

Bq **irazi** 'strain, filter' : Abkh **-ra-ʒa**, etc. < PAA ***-ǯə-**, Ubykh **za-**, Adyghe-Kabardian **zə-n**, PND ***(ɦ)ɔrǯ̌V-**, all 'strain, filter' (C85 [30])

Exactly the same parallel with Abkhaz is noted by Bouda (1948:§36). As so often, however, the Caucasian forms do not appear to be long enough to offer a convincing matchup with the Basque verb, whose root is *-raz-*.

[169.] Bq **isuri** 'pour (out)'

Bq **isuri ~ ixuri** 'flow' : PNC ***šorV** 'lake, river' : Sum **šur** 'flow' (B9lc PS2 [35])

Basque *isuri* doesn't mean 'flow'; it's a transitive verb, and it means 'pour, pour out' (a liquid or a mass of small objects).

[170.] Bq **itain** 'tick'

Bq i-**thañ** 'tick', **thiña** 'tick' : PNC ***t'ānhV** 'nit' (B9lb [36]) (B9lc [19])

Basque *it(h)ain* is one of a cluster of forms meaning either 'tick' or 'leech', depending on locality; the others are *ikhain*, *akaiñ*, *lakaiñ*, *lakhoin*, *lithoin*, *izai(ñ)*, and *zizain*. In reviewing this collection, Michelena (1977:292) concludes that we have here extensive contamination between two unrelated words, and he suggests that the original words in question are most likely *akain* 'tick' and *izain* 'leech', neither of which is of any assistance to Bengtson. Moreover, *akain ~ lakaiñ* itself is a loan word from Romance, most likely from Gascon *lagagno*; see Corominas and Pascual (1980) under *legana*. The unrelated **tiña** is not 'tick', but 'ringworm', and it is a transparent loan from Romance: cf. Spanish *tiña,* French *teigne,* both from Latin TINEA 'ringworm'.

[171.] Bq **itoi** 'drop'

Bq i-**t(h)oi 'a** drop', **thu-** 'spit', i-**thoi**-tz 'roof gutter', i-**thu-**rri 'spring, fountain' : PNC ***tujV** 'spittle, spit' : Bur **thuu-**š 'a drop', **thi-** 'pour', **thuu-** 'spit' : Sum **tù** 'pour out, libate' (B9la [82]) (B9lc PS1) (B93) (R94b)

Basque has a large number of local words for 'drop', with *itoi* being confined to a part of the French Basque Country. The several words for 'roof gutter', such as *itogin ~ itoin ~ itoite ~ itoitz ~ itoki ~ itokin*, are apparently all derived from the verb **itogi ~ *itoki*, a variant of *idoki* 'remove', and hence have nothing to do with *itoi*. There is no reason even to suspect a connection between *iturri* and any of the other words, and, as for *tu*, see **tu** 'spit' below.

[172.] Bq **itsaso** 'sea'

Bq i-**tsaso** 'sea' : PNC (NW) ***čə** 'salt' : Bur **šau** 'oversalted' : PY ***čV?** 'salt' (B9lb [18]) (B9lc [41]) (B9ld) (SR94 [225])

[173.] Bq itsu 'blind'

Bq **itsu ~ utsu ~ utsi** 'blind' : PNC ***Hö̲ć^{w}V** 'dark, blind' (B9lc [55])

The Basque word for 'blind' is *itsu*; the other forms cited by Bengtson do not exist. What he has in mind is the Zuberoan variant *ütsü*, showing the usual high-vowel assimilation of that dialect, and its definite form *ütsia*. These transparently secondary forms seem to have been invoked merely to get a rounded vowel into the first syllable, in order to obtain a better match with Caucasian.

[174.] Bq **itzal** 'shade, shadow'

Bq i-**tzal** 'shade, shadow' : Bur **-yAl̲** ~ **-nAl** 'shade, shadow' (B91b [23])

[175.] Bq **izan** 'be'

Bq **iza**-n 'be' : PNC ***ʔosʷV** 'sit' : Bur **ōs** 'set' : Hatti **-š-** 'sit, set' : Urartean **aš-** 'sit; put' : PY ***hAs-** 'be' (B91b [61]) (B91c [62]) (SR94 [13])

The final *n* of *izan* is probably part of the stem; see Trask (1990).

[176.] Bq **izar** 'star'

Bq **izar** 'star' : PAA ***c'ʷa**, PND ***Hă-ʒʒwărʕi** (C85 [84])

Bq i-**zar**(r) 'star ': PNC ***Ha-ʒʒwarʕi**'star' : Bur lA-**čaar** 'morning star' (Venus) : Sum u(d)-**sar** 'new moon' (B91a [73]) (B9lc PS1, PS1') (B91d) (B93) (B94b [37]) (R94b)

Bouda (1949: Appendix §60) adduces some rather different resemblances, mostly involving Kartvelian languages.

[177.] Bq **izeba** 'aunt'

Bq **ize**-ba, **iza**-ba 'aunt' : PNC (NE) ***-ilćʷi** 'woman (relative)' : Hurrian **sali** 'daughter': Urartean **sAlâ** 'daughter' (B9lb [44]) (B91c [139])

[178.] Bq **izen** 'name'

Bq **izen** 'name' : Adyghe-Kabardian **c'a**, Ubykh p'-**c'a**, PND ***ʒʒwĕrʔi-** (C85 [59])

Bq i-**zen** 'name', B u-**zen** : PNC ***ʒʒʷĕr-ʔi**-'name' : Bur **sen**-As 'say, tell, call, name', **sen**-As 'named' : Sum **zi** 'name', **sa**$_4$ 'name, call by name' (B9la [137]) (B9lc [60]) (B91d) (B94b [9])

Somewhat unexpectedly, Bengtson (1994b:34) concludes that the Basque word is a loan from Berber *ism* 'name'. But the Berber word is itself a loan from Arabic *ism* (Jamal Ouhalla, p.c.), and hence cannot predate the eighth century AD.

[179.] Bq **izotz** 'frost, ice' : See under **hotz** 'cold'.

[180.] Bq **iztai** 'groin'

Bq **iztai** 'anus', **izte**-r ~ **iste**-r 'thigh' : WCauc ***ʂʷtV** 'genitals' : Bur **-Aṣ̌ç̌**-I ~ **-Aṣ̌ṭ**-iŋ 'small of the back, loins, reins, waist' (B9lb [12]) (B94a [23])

Bizkaian *iztai* is not 'anus', but 'groin'. Moreover, it is merely a compound of *iztar*, a Bizkaian variant of *izter* 'thigh', with the noun-forming suffix *-egi*, showing the usual loss of final *-r(r)* in composition (Michelena 1977:338).

[181.] Bq **izten** 'awl'

Bq **izten** 'awl' : Berber t-**isten**-t 'awl' (*t* feminine) (B94b)

Bengtson sees the Basque word as a loan from Berber.

[182.] Bq **izter** 'thigh': See **iztai** 'groin'.

[183.] Bq **izu** 'trembling', 'fear'

Bq **izu** 'fear' : Abkh **-sʷa-**ra (C85 [58])

Exactly the same parallel is cited by Bouda (1948:§43).

[184.] Bq **jaio** 'be born'

Bq **jaio** 'be born' : Abkh a-**i-**ra 'birth' (C85 [31])

The match between Basque *jaio* (< **e-aio)* and the Abkhaz root *-i-* is less than stunning. The same match is pointed out by Bouda (1948:§44).

[185.] Bq **jakin** 'know (a fact)'

Bq j-**aki-**n 'know' : PNC ***Hiq'V** / ***Hüq'V** 'know' : Bur **-hAki-**n 'learn' (B9lb [65])

[186.] Bq **jin** 'come'

Bq j-**i-**n 'come' : Abkh **-i-/ja-** (C85 [26])

The verb *jin,* confined to northern dialects, is phonologically anomalous, since *j* (< **e-*) is normally found only before *a* or *o*. This must therefore be a contraction of a longer form, and that longer form is not far to seek: it's the synonymous *jaugin* (Michelena 1977:516); *jin* and *jaugin* are in complementary distribution. This secure etymology destroys the match with Abkhaz, such as it was.

Similar parallels are cited by Bouda (1949:§72).

[187.] Bq **joan** 'go'

Bq j-**oan** 'go' : PEC ***ʔAʔwA-**n 'go' : PY ***hejAŋ** 'go' (SR94 [110])

[188.] Bq **josi** 'sew'

Bq **josi** 'sew' : PAA ***ś̮ʷV** 'weave', PND ***ʔörs̄V** 'weave' (C85 [25])

The Basque word is from earlier **e-os-*i.

Bouda (1948:§132; 1949: Appendix, §30) compares Basque *josi* with South Caucasian.

[189.] Bq **-k** (plural suffix)

Bq **-k** (noun plural suffix) : Abkh **-kʷa** (plural of "non-reasonable" class): Bur **-ko(ŋ)** (plural of certain nouns) (B94a [31])

Strictly, Basque *-k* does not pluralize nouns, but full noun phrases, and it is always phonologically bound to a preceding determiner: only the four definite determiners can be marked for number in Basque. Moreover, *-k*

appears only in the absolutive case, plurality being marked in all oblique cases by *-e-*. This leads most vasconists to believe that the plural suffix was originally **-g(e)*.

[190.] Bq **-k** (ergative case-suffix)

Bq-**k** (ergative), **-ik** (ablative-partitive) : PNC ***-k'V** (ablative, instrumental, direction, location, partitive) : Bur **-Ak** / **-ek** (instrumental) (B91d) (B93)

There is no reason to suppose that Basque ergative *-k* is connected with partitive *-ik*; on this last, see **-ik** (partitive).

[191.] Bq **-k** (second-person singular [intimate male] agreement marker in verbs)

Bq **-k** 'thou' : PNC ***Gu** 'thou' : Bur **gu-** 'thou' : PY ***kV-** 'thou' (R94b)

Basque *-k* is not a pronoun; it is an agreement marker in finite verbs. For discussion, see **hi** 'you' above.

[192.] Bq **kaiku** 'wooden cup or bowl'

Bq **kaiku** 'wooden bowl' : PNEC ***qwăqwV** 'basket, vessel' : PY ***qVk** 'spoon, scoop' (B91b [52])

The Basque word, with its initial *k-*, can hardly be native, and in fact it is an obvious loan from Latin CAUCU 'drinking vessel' (of Greek origin), with dissimilation of **au-u* to *ai-u* (Michelena 1957b:19, 1973:112, 1974:190, 1977:91).

[193.] Bq **kako** 'hook'

Bq **khako** ~ **krako** 'hook' : Georgian **k'ak'vi** 'hook' (B94b)

Very implausibly, Bengtson sees the Basque word as a loan from Kartvelian. Northern varieties of Basque use *gako* ~ *kako* ~ *krako* indifferently for both 'key' and 'hook'; southern varieties have specialized *gako* for 'key' and *kako* for 'hook'. The word is almost certainly of Romance origin, though most Romance words for 'hook' contain a nasal, as in Castilian *gancho*. The French Basque variant *krako*, with its initial cluster, indubitably shows the influence of French *croc* 'hook', of Germanic origin and cognate with English *crook*.

[194.] Aquitanian ****kala**

Basque-Aquitanian **kala** 'castle', (?) Bq (Z) **kaloia** 'hut, cabin' : PNC ***qəlV** 'house' (B91b [48]) (B91c [80])

The introduction of an alleged Aquitanian **kala* 'castle' is sheer fantasy: no such word is attested in Aquitanian with this or any other meaning, and Aquitanian, of course, doesn't even have any stems beginning with *k*. Bengtson explains that he has obtained this item by assuming that the city name *Calagurris* (in Roman times; modern *Calahorra*) is Aquitanian/Basque and by analyzing it as consisting of his **kala* plus Basque *gorri* 'red'. I know of no reputable scholar who would support such an analysis: the evidence strongly suggests that Aquitanian was not even spoken as far south as Calahorra in Roman times, and most scholars would opt for a Celtic or an Iberian origin for the name of this Celtiberian city; see Corominas and Pascual (1980) under *Calahorra*. As for *kaloi(a)*, this is nothing but a variant of *kaiola* 'cage', 'cosy little place', a loan from Romance (cf. Bearnais *cayole* 'cage').

[195.] Bq **kankano** 'large, clumsy person'

Bq (B) **kankano** 'large fruitstone, kernel, almond' : WCauc ***k'ank'a**, ECauc ***k'erk'enV** 'egg' : Bur **kaka:yo** '(walnut) kernel' (B94a [29])

There are several problems here. First, Basque *kankan ~ kankano ~ kankanu* has a range of regional meanings, including at least 'large, clumsy person', 'idle, lazy', and 'person who begs alms unnecessarily'. The sense cited by Bengtson is in fact a very severely localized one, recorded by Azkue only in a single small area. Second, this word, with its initial *k-* and its medial *-nk-*, cannot conceivably be of any antiquity in Basque. Third, a stem *cancan-* is very widely found in western Romance in a variety of "expressive" formations, including, for example, Castilian *cáncano* 'louse' and Old French *cancan* 'loud noise' (which is apparently the source of the name of the well-known dance). In view of this, Corominas and Pascual (1980) do not hesitate to derive the Basque word from a Romance source.

[196.] Bq **kasko** 'skull', 'crown'

Bq **kasko** 'head, skull, summit' : PNC ***k̄ʷi[l]kʷV** 'skull, back of the head' (B9lb [l])

With its initial *k*, Basque *kasko* cannot possibly be a native word. In fact, its source could hardly be more obvious: it is Castilian *casco* 'skull', 'helmet'. This Spanish word has been borrowed throughout western Romance; see Corominas and Pascual (1980).

[197.] Bq **karats** 'bitter'

Bq **kharr**-atx ~ **garr**-atz 'bitter, sour' : PNC ***q'(q')elʔV** ~ ***q'(q')eq'elʔV** 'bitter' : Bur **ɣaaqAy** ~ **ɣAqaauy** 'bitter' : PY ***qʌqʌr** 'bitter, bile' (B9la [118]) (SR94 [21])

This strictly eastern word is *karats*, not **k(h)arratx*, and it means both 'bitter' and 'foul-smelling, fetid'. With its initial *k-*, it cannot possibly be of any antiquity; it is surely a loan word, though I haven't yet been able to identify a source.

[198.] Bq **kokot** 'nape'

Bq **kokot** 'nape' : PY ***kəqənt-** 'neck' (SR94 [179])

Basque *kokot* is not remotely native: it is a transparent loan from the Romance word which appears as Castilian *cogote*, Old Spanish *cocote,* Occitan *cogòt*, and so on, all 'nape'. This word is found throughout much of the Iberian Peninsula, throughout Occitan, and in the Catalan of Roussillon; it is regarded by virtually all romanists as a derivative of the widespread Romance word *coco ~ coca* 'round thing, head', which is regarded as an "expressive" formation, possibly a nursery word, and which is unattested in Basque. See Corominas and Pascual (1980), and see also **garkotx(e)** 'nape', above.

[199.] Bq **korotz** 'dung'

Bq **khorotz ~ gorotz** 'dung' : PNC ***k'urč'V** 'dung' : Bur **ɣurAš** 'dung' (B9lb [16]) (B9lc [53], [94])

The word *k(h)orotz ~ gorotz* is one of several Basque words for 'dung' or 'manure'. The widespread occurrence of the form with initial *k* strongly suggests that the word is not ancient: original **korotz* would have been impossible, and original **gorotz* should not have undergone devoicing. Corominas and Pascual (1980; see under *coroza)* conclude that the Basque word is derived from a Romance development of Latin CROCEA 'saffron-colored', which is phonologically impeccable.

[200.] Bq **kuma** 'mane, horsehair'

Bq **khuma** 'mane' : PNC ***q'(q')amhā** 'long hair, mane' : Bur **qAma** 'felt' (B9lb [4]) (B9lc [105])

Basque *k(h)uma* 'horsehair, mane' is a simple borrowing from Old Spanish *coma* 'horsehair. mane', itself from Latin COMA 'hair'.

[201.] Bq **kurlo** 'crane'

Bq **khur**-lo 'crane' : PNC ***q'(q')ɨrɨ**-q'(q'^{w})V 'crane' : Bur **qArū-ro** 'heron' (B9la [96]) (B9lc [103]) (B9le)

Basque *k(h)urlo* is a variant of *kurrillo*; as its initial *k* suggests, this is a loan from Romance, and it comes from the same Romance source as Spanish *grulla* 'crane'. See the Spanish word in Corominas and Pascual (1980) for discussion.

[202.] Bq **-la** (suffix forming adverbs of manner)

Bq **-la** (suffix of manner of activity) : Abkh **-la** (instrument and manner) (C85 [2])

The Basque suffix *-la* forms a handful of adverbs of manner, such as *hala* 'in that way' and *nola* 'how?'. It is possibly related to the suffix of identical form which serves as the [- WH] complementizer of the language. Exactly the same resemblance is pointed out by Dumézil (1933:128-129) and, at some length, by Lafon (1951:§7). It is not obvious that the Abkhaz morph has any Caucasian standing.

[203.] Bq **lagun** 'companion, friend'

Bq **lag**-un 'friend, companion' : PND ***ɬăGV̆** 'slave' (C85 [65])

Once again C is removing a morph he can't match and declaring it a ''suffix'', with no shred of justification. The semantic matchup proposed by C is also somewhat unexpected. Bouda (1948:§4) draws some rather different parallels between *lagun* and some Caucasian words.

[204.] Bq **larre** 'pasture'

Bq **larre** 'pastureland : Svan **lare** 'meadow' (B94b)

Bengtson sees the Basque word as a loan from Kartvelian.

[205.] Bq **laster** 'quick, soon'

Bq **la(i)ster** 'quick', 'soon' : Abkh **lassə**, a-**las** 'quick', Ubykh **was** 'quick', Proto-Lezgian ***s̄ɨlä** 'light', with metathesis (C85 [22])

The original form is clearly *laster* (Michelena 1977:159). But the putative Caucasian cognates leave the Basque element *-ter* completely unaccounted for. There is no evidence within Basque for a suffix of the form *-ter*.

Bouda (1948:§53) draws exactly the same parallels as C, though Bouda at least realizes the need to account somehow for the morph *-ter*.

[206.] Bq **lau** 'four'

Bq **lau** 'four' : PAA *p-**k̆'**ə (C85 [20])
Bq **lau**-, **lau**-r 'four' : Sum **limmu**, **lim** 'four' (B9lc PS2 [27])

It is clear that the Basque numeral was originally *laur*: this is still the form in the northern dialects, and all dialects have plural *laurak* 'all four', with the article *-a* and plural *-k*. This fact is of no assistance to C or B.

[207.] Bq **lo** 'sleep' (n.)

Bq **lo** 'sleep' : Sum **lu** 'lie' (B9lc PS2 [47])

[208.] Bq **lorratz** 'track, scent'

Bq ***lor-** in B, G *lorrats* 'track, trail, trace, scent' : WCauc ***ɬa** 'foot', ECauc ***leɬ̄hV** 'sole' (B94a [27])

The Basque word is not **lorrats* but *lorratz*. The second element is the widespread *atz* 'trace, vestige', which in the cited word has been reinforced by the B G noun *lor* 'log which is dragged', and also 'act of dragging'; the word means literally 'track left by a log which is dragged', but it is commonly used in a figurative sense. There is no connection with 'foot'.

[209.] Bq **magal** 'lap'

Bq **magal/bagal** 'belly' : Abkh **a-mgʷa**, **a-bgʷ a**, Ubykh **nək'ʷə** 'belly' (C85 [45])

The Basque word *magal* does not mean 'belly', which is *sabel*. Rather, it means 'lap' and, by extension, 'protection', 'refuge'.

[210.] Bq **maguri** 'strawberry'

Bq **maguri** 'strawberry' : PAA ***maRa** 'blackberry bush', PND ***mäwq̄'V** 'strawberry' (C85 [78])

The Basque word for 'strawberry' exhibits an extraordinary variety of regional forms, most of them of the general pattern *ma(u)XuYi*, where *X* is any of *r*, *rr*, *l*, or *ll*, and *Y* is any of *g*, *b*, *k*, *h*, or zero, or else the other way round. The most widespread form is *marrubi*, but many others exist, including *mallubi*, *marubi*, *mahurri*, *magauri*, *maguri*, *malubi*, *malhuri*, *mailluki*, *maillugai*, *mauli*, *mauliki*, *maulubi*, and so on. It is impossible to determine the earliest form of this word, but forms like C's *maguri*, with a velar consonant at the beginning of the second syllable, are confined to a very small area in High Navarre; the far more widespread forms like *marrubi* and *mallubi* surely represent the ancestral form better, but are of no assistance to C. In any case, the word-initial *m-*, as always, cannot be original in Basque: it must derive from pre-Basque **b-*, further destroying the proposed parallels.

[211.] Bq **makutsik** 'in one's shirtsleeves (adv.)

Bq **-kuts-** 'sleeve' in ma-**kuts**-ik 'en mangas de camisa' : PNC ***q'q'ʷV̄Ič'ʷV** 'sleeve, elbow, armpit': Bur **quš** 'elbow of a garment' (B9lb [53])

A word *makutsik* 'in one's shirt-sleeves' is preposterously analyzed as containing a root **-kuts-* 'sleeve'; in fact, this is a simple compound of *mauka* 'sleeve' (a loan from Latin MANICA 'sleeve') plus *(h)uts* 'empty, plain, bare' and partitive *-ik*, which derives adverbs from N-bars (compare *eskutsik* 'bare-handed', from *esku* 'hand' and *[h]uts*). The loss of final *-a* in the first element before a following vowel is absolutely regular, and the dissimilation of *au* to *a* before a following *u* has ample precedent in Basque. Bengtson's putative **-kuts-* is another fine example of the dangers of performing comparisons by extracting items *ad libitum* from dictionaries and segmenting them at whim.

[212.] Bq **mama** 'liquid' (nursery word)

Bq **mama** 'water, potable liquid' (child speech) : Bur **mamu** 'milk, sap (of crops)' (B94b [17])

It strikes me as absurd to pursue remote cognates for what is transparently a nursery word of no great antiquity.

[213.] Bq **mama** 'breast', 'mother' (nursery word)

Bq **mama** 'breast; mother', **mama**-tu 'suckle' : PNC ***māmV** '(female) breast' : Bur **-maam**-ut '(female) breast, nipple (male or female), teat'; cf. **mamu** ~ **mAma** ~ **maama** 'mother, maternal aunt' (B9la [42])

Insofar as it exists at all, the rare French Basque *mama* is strictly a nursery word, and there is no reason to suppose it is anything but a loan from Spanish or French, both of which have similar words in these meanings. In any case, the existence of a form like "*mama*" in the sense of 'mother' or 'breast' is so widespread in the languages of the world, and so readily explicable in terms of babbling, that such forms cannot possibly serve as a basis for remote comparisons.

[214.] Bq ****maño** 'masculine'

Bq **maño** 'masculine, macho' : PNC ***mVnXV** 'man, male' : Bur **mẹn** ~ **mẹn**-ik 'people, some people, any people, someone, anyone; who?' : PY ***pix-** 'man': Sum **mu**, **mw** 'man' (B9lb [37])

An alleged **maño* 'masculine' is unknown to me and to the lexicographers, and could not be ancient if it did exist.

[215.] Bq **mara-mara** 'smoothly'

Bq **mara-mara** 'abundant' : Abkh i-**marə-maža** 'abundance', etc. (C85 [23])

There are several problems here. First, the Basque word does not mean 'abundant' and it is not even an adjective. It is an adverb, and Azkue's 1905 dictionary glosses it as 'smoothly', though his examples, involving the flow of sweat, tears and blood, suggest that a better gloss might be 'steadily, continuously'. Aulestia's 1989 dictionary glosses it as expressing 'silent but continuous movement' and illustrates it with reference to falling snow. Second, a phoneme */m/* cannot be reconstructed for pre-Basque, which strongly suggests that the creation of *mara-mara* in Basque postdates the (probably post-Roman) introduction of */m/* into the language. Third, this is a phonesthetic item, and phonesthetic items are notoriously treacherous objects for comparison. Modern Basque is particularly fond of */m/* in onomatopoeic and phonesthetic items: *murmur* 'murmur', *marmar* 'growl, grumble, meow', *zurrumurru* 'whisper, murmur, gossip', *murdurika* 'muttering, murmuring', *zirimiri* 'drizzle', *momorro* 'creepy-crawly', *marrao* 'yowl, meow', *marruma* 'roar', *maillo-maillo* 'slowly, gradually', and many others. Apart from the semantic problem, C's proposal would apparently require us to believe that pre-Basque exceptionally permitted */m/* in such items, in defiance of the ordinary phonology of the language. Such cases are by no means unknown: several languages are attested in which ideophones exhibit segments which are otherwise non-existent. But C nonetheless appears to be expecting us to swallow quite a bit in accepting this matchup.

[216.] Spanish **marrano** 'pig'

Spanish **marra**-no 'pig' : PNC ***mīrğV**, ***mirλ̆V** 'male, man' (B9lb [38])

I confess I have little idea what the point of this comparison is supposed to be. The Spanish word is one of several Romance words in *marr-* denoting male animals, and presumably Bengtson would have us believe that these are loans from Basque. But a Basque origin is out of the question, and is not even considered by Corominas and Pascual (1980) in their examination of these words.

[217.] Bq **matel** 'cheek', **moto** 'headscarf'

Bq **mat**-el 'cheek', **mot(h)**-o 'hairknot', etc. : PNC ***mət'V** 'face' : Bur **mAto** 'brains' : PY ***bat(t)-** 'face' (B91a [9]) (B94b)

In spite of his proposed matchups, Bengtson declares (1994b:34) that the Basque word *matel* is a loan from Berber *a-ma:ḍəl*. Now the Basque word for 'cheek' is variously *maxela ~ matel(a) ~ matrailla*, and Michelena (1977:188) derives it from Latin MAXELLA 'jaw'. The French Basque word *moto* cannot possibly be related; it most commonly means 'headscarf', but also 'child's bonnet', 'plait, braid (of hair)', 'bird's crest' or 'hairknot'. This too is doubtless a loan word, though I have not so far managed to identify the source. (One possibility for at least some of the senses is Occitan *moto ~ mouto,* which means both 'clod, lump, ball' and 'mound, elevation'.)

[218.] Bq **mihi** 'tongue'

Bq **mihi** 'tongue', **min**-tzo 'speech, language', **mili-** 'lick' : PNC ***mə̄lc'i** 'tongue' : Bur **-melč** 'jaw' : Sum **me**, e-**me** 'tongue, speech' (B91a [1]) (B91c [11]) (B91d) (B94a [7]) (B94b [8]) (R94b)

Once again: no ancient Basque word could begin with *m-*, because there was no *m* in the language. The comparative evidence for 'tongue' points indisputably to original **bini*, destroying the match (Michelena 1958:44; 1977:222). French Basque *mintzo* means 'voice, speech, conversation', and it is a transparent derivative of *mihi* showing the usual combining form *min* (< **bin-*). The Basque word for 'lick' is not **mili-* but *milikatu*, stem *milika-*. This has a variant *mihikatu*, suggesting that the word might be a derivative of *mihi*, but equally the word might have an entirely distinct origin (as I suspect), with *mihikatu* then being a folk etymology.

[219.] Bq **moto** 'headscarf': See under **matel** 'cheek'.

[220.] Bq **muin** 'pith, marrow, inner part'

Bq **mun ~ mu(i)ñ** 'brain' (< ***mVHV** + **hu(i)ñ** 'marrow') : ECauc ***maʕu** 'brain' : Sum **múh** 'top, skull' (B91c PS2 [51])
Bq **huñ** 'marrow' : PNC ***hʷenʔV** 'blood' : Bur **hAn** 'blood' (B91c [118]) (B91d)
Bq **mun** 'medulla', **munak** 'brains' : PW ****MENA** 'think about' (BR94 [16])

There is a great deal of confusion here. First, Basque *muin ~ muñ ~ mun ~ (h)un* are not two different words, as Bengtson seems to think: they are only regional variants of a single word. Second, the word means 'pith, marrow, inner part'. It forms a compound *burumuin* 'brain' (*buru* 'head'), which on occasion is reduced to *muin*, but the sense of 'brain(s)' is clearly derivative. Third, the regional variation points strongly to an original variation **bune ~ (h)une*, which is readily explained if this is a loan word from a Latin *FUNE (cf. *biku ~ iko* 'fig', from Latin FICU). Interestingly, Schuchardt proposed a century ago that the Basque word was borrowed from Latin FUNE 'rope', which is phonologically perfect if semantically unexpected; Michelena (1977:150) accepts the value of this proposal without embracing it wholeheartedly.

[221.] Bq **mulo** 'pile, stack (of hay or cornshocks)'

Bq **mulho** 'small hill' : PNC ***muʕalV** 'mountain' (B93)

The Basque word means 'stack' (commonly 'haystack'), and it is a transparent loan from Romance: cf. Occitan *moulou ~ mouloun*, Fr *meule ~ meulon*, of identical meaning, derived by Meyer-Lübke (1935) from a Gallo-Romance *MŪLA. The sense of 'small hill' is a hapax cited by the Dutch linguist van Eys; it has no other attestation, and is not the sort of thing that should be cited unblushingly as the only sense of the word.

[222.] Bq **musu** 'kiss', 'face', **musin** 'snout, muzzle'

Bq **musu** 'lip, face, mouth, snout, kiss', **mus**-in 'snout, muzzle' : PNC ***mɔč'V** 'braid, hair, beard' : Bur **muš**-k-Ane 'on one's face, face down', **muš**-ki 'pubic hair' ('the beard below'), PY ***bunč-** 'face, chin' : Sum **muš** 'face, appearance' (B9la [16]) (B9le)

The Basque words are transparently of Romance origin: cf. Old Spanish *muso*, Old French *muse*, *musel*, French *museau*, Occitan *mus*, *museu*, all 'snout, muzzle' and all from late Latin MŪSU or from a diminutive of this (English *muzzle* derives from the same source).

[223.] Bq **mutur** 'snout, muzzle'

Bq **mut(h)ur** 'snout, muzzle' PNC ***mərǎ'V** 'nose, beak' : Bur **-multur** 'nostril' (B9la [23]) (B9lc [12], [77])

The word *mutur* is almost certainly of Romance origin; see Corominas and Pascual (1980) under *modorro*.

[224.] Bq **-n** (locative case-suffix)

Bq **-n** (locative suffix) : Abkh **-n**, PND ***-nɨ** (locative) (C85 [8])
Bq **-n** (inessive) : PNC ***-nV** (genitive): Bur **-Ane/-Aŋe/-eŋe** (comitative) (B9ld)

This matchup, while seemingly straightforward enough, runs into an unusual problem. An element *n* with a locative sense is found in a staggeringly large number of the world's language families. A locative affix *-n*, *-Vn*, or *-nV*, or a locative particle *Vn* or *nV*, is found, for example, in Indo-European, Uralic, Dravidian, Kartvelian, Afro-Asiatic, Altaic, and Japanese (at least), where it forms one of the pieces of evidence adduced in favor of the Nostratic super-family (Kaiser and Shevoroshkin 1988:313; Shevoroshkin and Manaster Ramer 1991:181); a similar element occurs in many languages of North and South America, where it is cited by Greenberg (1987:302) as evidence for his Amerind super-family (and Hymes [1987:661] adds more *-n* locatives overlooked by Greenberg); and such examples could be multiplied *ad libitum*. This *n* is precisely the kind of thing singled out by the proponents of "Proto-World" as evidence for monogenesis. Of course, such widespread occurrence does not, of itself, prevent C from citing it here, but, assuming we are not enthusiastic about Proto-World, it seems that we may here be looking at one of those quasi-universals (like the worldwide occurrence of *m* in first-person singular forms and in words for 'mother') whose explanation is to be sought outside the confines of a genetic relationship.

As for the second matchup, it amounts to nothing more than the banal observation that lots of case languages have some case-ending or other with an *n* in it.

[225.] Bq **-n** (past-tense suffix)

Bq **-n** (past-tense marker) : Abkh **-n** (past tense), PND ***-na** (gerundial past tense suffix) (C85 [7])

The match with Abkhaz has been pointed out many times before, for example by Dumézil (1933:146), by Lafon (1944:532-533) and by Lafon (1952a:§28). It is not obvious that a past-tense *-n* has any Caucasian standing.

[226.] Bq **negu** 'winter'

Bq **negu** 'winter' : PNC ***Hǎ'^{w}i̲nī** 'winter' (metathesis) (B9lc [75])

The match with Basque *negu,* such as it is, is obtained only by invoking a metathesis.

[227.] Bq **ni** 'I'

Bq **ni** 'I' : PND ***nɨ̌*** 'I' (C85 [11])
Bq **ni** ~ **ni-** ~ **ne-** 'I, me' : PNC ***nɨ*** 'I' : Sum (?) **ní** 'self' (B9la [138]) (B9ld) (B94b [l])

At first glance, the Basque-Caucasian parallel is striking. A difficulty, however, is the very great frequency of nasals in first-singular pronouns in the world's languages. Even leaving aside those in *m* (as in Indo-European, Uralic, Altaic, and Kartvelian), first-singular pronouns with coronal or velar nasals are abundantly attested in Afro-Asiatic, Athabaskan, Australian, Dravidian, Eskimo-Aleut, Muskogean, and Niger-Kordofanian, just to name a few examples. Therefore, given the apparent absence of Burushaski and Yeniseian forms with nasals, it appears that little weight can be assigned to the Basque-Caucasian match. This is particularly so for the following reason. Personal agreement markers in Basque generally correlate strongly with the corresponding pronouns: e.g., *gu* 'we', agreement markers *g-* (prefix) and *-gu* (suffix). For *ni,* the markers are *n-* (prefix) but *-da* (suffix; see under **-t** below). This suggests an ancient stem-alternation in the Basque pronoun, and once again (as with **hi** 'you' above) this is just the sort of puzzle on which we would expect a valid comparison to shed some light — but no such light is forthcoming here.

The Caucasian matchup has been pointed out several times before, for example, by Dumézil (1933:138), by Lafon (1944:383, 390) and by Lafon (1952a:§22).

[228.] Bq **nigar** 'tear' (n.)

Bq **niga**-r ~ **nega**-r 'tear' : PNC ***newq'u*** 'tear, pus' : Bur **nAgei** 'boil' : Sum **mag**-(bi), **mak**-(kaš) 'lamentation' (B9la [6]) (B9lc [29]) (B94b [16])

[229.] Bq **nor** 'who?'

Bq **no-** (interrogative particle) : Abkh **-n** (particle in interrogative auxiliaries), Proto-Lezgian ***ni-*** (interrogative pronoun) (C85 [10])
Bq **no**-r ~ **nu**-r 'who', **no**-iz ~ **nu**-iz 'when' : PNC ***nV*** (interrogative) : Bur (W) a-**na** 'where, whither', a-**n**-Um 'whence' : PY ***?an-*** 'who' : Sum a-**na** ~ **en** ~ **en-na** 'what', ma-**na** ~ me-**na** 'when' (B9la [147]) (B9lc [30]) (B9ld) (SR94 [295]) (B94a [6]) (B94b [7]) (R94b)

Basque *no-* is not well described as an "interrogative particle". It is rather one of the two interrogative stems of the language, the other being *ze-*. The first appears in such forms as *nor* 'who?', *non* 'where?', *noiz* 'when?' and *nola* 'how?', the second in *zer* 'what?' and *zelan* 'how?', among others. In any case, the matchup with Caucasian has been pointed out many times before, for example, by Uhlenbeck (1924:578), by Trombetti (1925:§115), by Dumézil (1933:61, 141), by Lafon (1944:535), and by Lafon (1951:§ 7). In the last work (§18), Lafon in fact points out that it is Basque *ze-* which finds the more impressive matchup in Caucasian languages, all groups of which (North and South) appear to possess interrogatives in *s-*.

[230.] Bq **odol** 'blood'

Bq o-**dol** 'blood' : Bur **del** 'oil; contents of an egg' : Sum **dal** 'breath, life, soul' (B91a [47]) (B9lc [25]) (B9ld PS1) (B93) (R94b)

[231.] Bq **ohe** 'bed'

Bq **oe** (**oi**) 'bed' : Abkh a-**ja**-ra, a-la-**ja**-ra 'lie down', a-**ja**-r-ta 'bed' (C85 [61])

Bouda (1949:§78) draws a rather different parallel between the Basque word and the Cherkess word for 'nest'. It is very difficult to see that the rather formless Basque word *ohe* can be of much assistance in a remote comparison of the sort being undertaken by C.

[232.] Bq **ohoin** 'thief'

Bq o-**hoin** 'thief' : Bur **γēn** 'thief' (B9lc [127])

As in so many other cases, Basque *ohoin* shows clear evidence of a lost nasal, and Michelena (1977:140) reconstructs a stem in **ono-*, which destroys the match, such as it was.

[233.] Bq **oihan** 'forest'

Bq oi-**han** 'woods, forest' : PNC ***fanV** 'mountain, shady slope of a mountain' : Bur **hūn** 'wood, timber' : Hurrian **fâvana** 'mountain' : Urartean **vâvânV** 'mountain' (B9lb [25]) (B93)

As always, Basque *h* is not etymological, and the arbitrary segmentation is particularly severe in this case.

[234.] Bq **oin** 'foot'

Bq **hoin** 'foot' : PNC ***ʔʷin**-qʷV 'heel' : Bur **-γān** 'heel' (B9lc [125]) (B9ld)

The word for 'foot' is *oin,* Zuberoan *huin*; Bengtson's alleged *hoin* is completely unknown to me, and the *h* is, as always, unetymological.

[235.] Bq **on** 'good'

Bq **hun** ~ **hon** ~ **un** ~ **on** 'good', **hun**-tu 'improve, ripen' : PNC ***hʷin**-žV ~ ***hIin**-žV 'good', ***Hün**-tV ~ ***Hin**-tV 'love, want', ***Hin**-ğğʷV 'love, want' (B9la [116]) (B9lc [117]) (B9ld)

The Basque word for 'good' is universally *on*, except in Z, which has *hun*, showing the usual aspiration of monosyllables in that variety as well as the regular raising of *o* to *u* in final syllables. So far as I am aware, neither **hon* nor **un* exists anywhere. The Z verb *huntu* is a simple derivative of *hun* (other dialects have *ondu*) and should not be cited separately. The Aquitanian evidence suggests that the earliest form of the adjective might have been **bon,* with the frequent loss of initial *b* before *o*.

[236.] Bq **or** 'dog'

Bq **xor** 'dog' (Z) : PAA ***xwa**, PND ***xwar** 'dog' (C85 [72a])
Bq **hor** ~ **or** ~ /xor/ (Z) 'dog' : PNC ***Xʷělje** 'dog', (NE) **xʷar** 'dog' : Sum **ur** 'dog' (B9la [87]) (B9lc [97])

The word for 'dog' is normally *or*, even in the aspirating dialects; *hor* is found only in Zuberoan, the dialect which regularly extends the aspiration to vowel-initial monosyllables.

C's alleged *xor* is presumably supposed to be a form of this *(h)or*, but, once again, the form is nowhere attested and appears not to exist. If it does exist, it can only be another expressive form, with the palatal consonant *x* prefixed for this purpose. Such a procedure for constructing expressive forms is all but unattested in Basque: Michelena (1977:189) can cite only three parallels (*txonil* < *onil* 'funnel', *txaul* < *ahul* 'weak', *txingura* ~ *xinguri* < *ingude* 'anvil'). (More usually, an expressive form is derived by replacing another consonant with a palatal consonant.) Even if C's form does exist, then (and I suspect that it does not), it cannot represent an original *x*-, and the already faint match with Caucasian is gone.

[237.] Bq **ortzi** 'sky, thunder, storm'

Bq **hortz** 'cloud' : Bur **xurōnč** 'cloud' (B9lc [113])
Bq **oz** 'sky', **Ost**-egun 'Thursday' : PNC ***ʔamsV** 'sky, sky-god' : Bur **aiyAš** 'sky' (B9lb [19]) (B9lc [124])

There is some confusion here: Bengtson is citing the same Basque word twice and comparing it with two different Burushaski words. The word in question is *ortzi*, which, occasionally alone but more usually in compounds, means 'sky', 'thunder' or 'storm', or, very rarely, 'cloud'. Its combining form is *o(r)(t)z-*, and it is the combining form *oz-* which Bengtson has wrongly cited as though it were an independent and different word. It is probable, though not certain, that *ortzi* is present in the name of Thursday: eastern *ortzegun*, western *ostegun*, with the common western development *-rtz-* > *-st-*. See Michelena (1971c) or Trask (forthcoming c) for a discussion of this.

[238.] Bq **osin** 'deep place in a river'

Bq **hosin** 'well' : PNC ***ʔʷɨn̲(c)cV** 'well' : PY ***sin-** 'spring' (B9lb [21]) (B9lc [35], [126])

Basque *osin* (usually so, even in the aspirating dialects) does not mean 'well', but rather 'deepest point in a river'; the confusion arises because Spanish *pozo*, correctly used to gloss the Basque word, has both meanings. Moreover, Bengtson is obliged to posit a metathesis to make this match work.

[239.] Bq **oskol** 'shell, peel, bark (of a tree)'

Bq (B) **oskol** 'nail, claw', dialect variants **ezkazal, azazkal, azüskülü** (no cognates offered) (B94b [14])

Bengtson (1994b:34) concludes that the Basque word is a loan from Berber *isker* 'nail'. But Basque *oskol* means 'shell, peel'; its compound *atzoskol* (*[h]atz* 'finger') means 'fingernail', and this is sometimes reduced to *oskol*. Moreover, the word appears in many localities as *koskol*, and it is just one member of a large family of words (*koskor*, *kozkor*, *kozkur*, *koxko*, etc.), all meaning 'chestnut shell', 'hard bread crust' or 'hard outer covering' in general. Thus it can hardly be separated from the equally large group of similar words in Castilian, Aragonese, and Occitan, such as Castilian *coscurro* 'hard bread crust' and Occitan *couscoùlho* 'fir cone'. All these words, and many others, are derived by Corominas and Pascual (1980) from the onomatopoeic item **cosc-* 'blow to a hard object' (in fact, though Corominas and Pascual do not note this, the word *kosk* is attested in precisely this sense in northern varieties of Basque). The other words cited by Bengtson are not "dialect variants", but completely unrelated words; the last one doesn't even mean 'fingernail', but only 'hoof'. See under **azal** 'skin' above.

[240.] Bq **oso** 'whole, complete'; 'healthy'

Bq **oso** 'whole, complete' : PAA ***za** 'whole', PND ***ʔɔ̄c'V** 'be whole' (C85 [42])
Bq **oso** 'full, complete, whole' : ECauc ***ʔoc'V** 'be full, fill' : PY ***ʔute** 'full' (B9la [110]) (B9lc [123]) (B94a [19]) (B94b [31])

Bouda (1949:§9) also compares Basque *oso* with some North Caucasian words. The Basque word does not mean 'full'.

[241.] Bq **otso** 'wolf', **potzo** 'mastiff'

Bq **ots** 'wolf', **potzo** 'wolf' (Z) : PAT ***baʒa**, PND *He-**bĕrc'ĭ** (C85 [72])
Bq **otso** 'wolf' ~ (Z) **potzo** 'wolf, big dog' : PNC *Hĕ-**bĕrc'ĭ** : PY ***pes**-tap 'wolverine' (B9la [89]) (SR94 [304])

The Basque word for 'wolf' is *otso,* not **ots*; this word is probably present in Aquitanian *Oxson*, *Osson* (Michelena 1954a:434). Bouda (1948:§59) cites similar parallels with North Caucasian words. The Zuberoan word *potzo* does not mean 'wolf'; it means 'large dog, mastiff', and it cannot possibly be either ancient or related to *otso*, which exists in Zuberoan as elsewhere; on this, see Michelena (1953:143).

[242.] Bq **paru** 'stick of wood, pole'

Bq **phau** 'stake, pole' (if not from Romance **palu-**) : Bur **-pauwo** 'stick, staff' : Sum **pa** 'staff, sceptre' (B9lc PS2 [17])

Basque *phau* is nothing more than a Z form of *paru*, which most certainly is a loan from Latin PALU (Z has recently lost intervocalic *a*).

[243.] Bq **pataxa** 'bottle'

Bq **phata**-sa 'bottle' : PNC ***pat'V** 'a kind of vessel' : Bur **pfAta** ~ **pfAte** 'wooden dish or bowl' (B9lb [51]) (B91c [3], [24]) (B9ld)

The word is not **phatasa* but *pataxa*, and it is attested only in a *single* small area in Navarre. It may perhaps be related to the widespread Ibero-Romance (and Occitan) *patax ~ pataxa ~ pataxe ~ patatxo ~ patache* (of Arabic origin), denoting a type of (sailing) vessel, but it is certainly neither native nor ancient.

[244.] Bq **pinpirin** 'butterfly'

Bq **pin-pir**-in ~ **pin-pil**-in 'butterfly' (reduplication and dissimilation) : PNC ***pɔ̄rV** ~ * **pɔ̄lV** 'bee, butterfly' (cf. ***purV** 'fly' [v.])) : Bur **phir**-An ~ **pfIr**-An 'spider; soul' ~ **pfer**-An 'moth; soul' : PW ****PAR** 'fly' (B9la [97]) (BR94 [19])

The citation of Basque *pinpirin* 'butterfly' as representing PW **par* 'fly' I find extraordinary. To begin with, this is merely one of a range of quite different regional words for 'butterfly', and one confined to the Lapurdian dialect of the French Basque Country. Further, I must point out yet again that *no* native Basque word of any period begins with *p*, except for phonesthetic items. And that's exactly what this word appears to be: compare the words *pinpa* 'bounce', *pinpoil* 'somersault' and also 'bubbling', *pinpi-panpa* (and also *panpa*) 'smack!' 'bangl', *panpina* 'child's doll', *panpalina* 'small bell', *panpotsa* 'throbbing' and *pinpili-panpala* 'favorite' (all from the same dialect) and *pinpin* (a certain finger game played by children) (from an adjoining dialect). It is clear that phonesthetic formations involving the morphs *pin-* and *pan-* are highly favored in this area, and there is little reason to regard the word for 'butterfly' as any different from the others: given the frequency of *pin-* in such formations, Bengtson and Ruhlen's proposed **pir-* ("with dissimilation") represents nothing but special pleading. Anyway, flight is hardly the most conspicuous characteristic of butterflies, and butterflies are hardly likely to be the creatures singled out above all others as the quintessential flying creatures.

[245.] Bq **pintza** 'membrane'

Bq **phintz(a)** 'membrane' : PNC ***pe̲n(c')c'ʷV** 'eyelash' : Bur **-phiinIšo** ~ **-pfīnišo** '(human) hair of the head' (B9la [14]) (B9lc [1], [48])

The word is more properly *pintza*, with no aspiration, and it denotes the membrane covering an egg or a nut. As I hardly need to remind the reader at this stage, no native Basque lexical item of any period begins with *p-*, and hence the Basque word cannot be native. In fact, it is an obvious loan from the Romance word which appears in Castilian as *binza* 'thin skin on the body of an animal' and in Aragonese Spanish as *binza* 'membrane covering an onion or an egg'. This word is of Latin origin; see Corominas and Pascual (1980) under *binza* for a discussion. The Basque *p-*

is admittedly unexpected, but, in recent loans from Romance, Basque often fails to respect the voicing of initial plosives: for example, in the Basque of Renteria, Spanish *corbata* 'necktie' and *gabarra* 'barge' are borrowed as *gorbata* and *kabarra,* respectively (Michelena 1957), and the Romance loan *berde* 'green' (cf. Castilian *verde,* French *vert*) appears in some localities as *perde.* The proposed semantic match would be far-fetched even if the Basque word were native.

[246.] Bq **pipil** 'bud'

Bq **pipil** 'bud' : PY ***bajbAl** 'bud' (SR94 [36])

Western *pipil* looks like anything but a native Basque word. Northern varieties have a verb *bip(h)ildu* 'pluck (a chicken)' and an associated word *bip(h)il,* which as an adjective means 'plucked, denuded' and as a noun means 'twig'. This group of words is derived by Michelena (1977:64, 258) from Latin DEPILARE. It is probable that western *pipil* is the same word with voicing assimilation; such voicing assimilation is sporadic but common in western varieties, and the forms *pipil* and *bipil* are in complementary distribution.

[247.] Bq **potorro** 'vulva'

Bq **photo-**rro 'vulva' : PNC ***pət'i** 'vulva' (B9lc [23])
Bq **poto-**rro 'pubis, vulva' : PW ****PUTI** 'vulva' (BR94 [21])

Basque *potorro* is just one of a large number of regional words for 'vulva', all formed on an element *pot-* or on a palatalized variant *pott-* or *potx-*; see Etxezarreta (1983) for a catalog. All these words are coarse slang terms, not in decent use. With their initial *p-*, these words cannot possibly be native. Similar words are attested throughout Iberian Romance and Occitan, and even Italian exhibits *potta* 'vulva'; hence the Romance origin of the Basque words can hardly be doubted. See Corominas and Pascual (1980) under various entries in *pot-* and *pach-*.

[248.] Bq **potzo** 'mastiff': See under **otso.**

[249.] Bq **ra-** (causative)

Bq **ar-** (causative prefix) : Abkh **ar-** (causative prefix) (C85 [1])

The ancient Basque causative prefix is not **ar-,* but *ra-*: *e-gin* 'do', *e-ra-gin* 'cause to do'; *i-bili* 'go about, move', *e-ra-bili* 'cause to move; use'; *joan* 'go' (< **e-oan), eroan ~ eraman (<* e-ra-oan)* 'take away'; *i-kusi* 'see', *e-ra-kutsi* 'show'; *i-kasi* 'learn', *i-ra-katsi* 'teach'; etc. This Basque-Abkhaz similarity is pointed out by Trombetti (1925:§121), by Dumézil (1933:149), by Lafon (1944:425-428) and by Lafon (1952a:80), though, according to Lafon, the Abkhaz morph is *r-,* not *ar-*: *c-* 'go', *r-c-* 'cause to go'. The Abkhaz morph appears to have no Caucasian standing.

[250.] Bq **-ra** (allative case-suffix)

Bq **-ra** (aditive) : PNC ***-rV** (locative) : Bur **-ar** (allative) (B9ld)

The Basque allative suffix is variously *-a* or *-ra.* Since, broadly speaking, *-a* occurs with pronouns, demonstratives and place names, while *-ra* occurs elsewhere, the most plausible interpretation is that *-a* is the original case-ending and that *-ra* has been introduced by reanalysis of forms like *hara* 'thither', in which the *r* is part of the stem. This destroys the matchups.

[251.] Bq **sabel** 'stomach, belly'

Bq **sab-el** 'stomach, abdomen' : PNC ***ǯabV** 'kidney, liver' : Bur **-A-so** 'kidney' : PY ***tVp-Vl** 'spleen' : Sum **šà.g** 'heart, breast, belly, guts' (B9lb [8]) (B9lc PS2 [4]) (B9ld)

[252.] Bq **sagar** 'apple'

Bq **sagar** 'apple' : (Cauc) Lezgi **čüX^{w}er** (and others) 'pear' : Bur **šuɣuri** ~ **šoɣori** 'a kind of pear' (B9lb [30]) (B9ld)

[253.] Bq **sagu** 'mouse'

Bq **sagu** 'mouse', 'rat' : PAA ***ccəɣ^{w}**-(a) 'mouse', 'marten', PND ***cwarğğwV** 'weasel', 'rat' (C85 [71])

Bq **sagu** ~ **sabu** 'rat, mouse' : PNC ***c^{w}a̲r(ğ)ğ^{w}V** 'weasel, marten, mouse' : Shina **čhArgeei** 'woolly flying squirrel' (presumably from Burushaski substratum) : PY ***saʔqa** 'squirrel' (B91a [92]) (B91c [34], [91]) (B9ld) (B9ld PS1) (R94b)

Basque *sagu* means only 'mouse', never 'rat', and the localized variant *sabu* is clearly secondary. The word for 'woolly flying squirrel' is puzzling: in several articles Bengtson cites this as "Burushaski", yet in his (1991a) article he describes it as "Shina — presumably from Burushaski substratum". I find this distressing, since we apparently don't even know what language is being cited. Bouda (1948:§66) also compares Basque *sagu* with North Caucasian words, but both he and Michelena (1967:607) point out that this Basque word has far more impressive parallels in the Kartvelian (South Caucasian) words for 'mouse'.

[254.] Bq **sahats** 'willow'

Bq **sahats** 'willow' : Bur **šAsk** 'willow' (B91c [64])

The word for 'willow' is variously *sahats* ~ *sagats* ~ *sarats* ~ *saats*; this type of variation points unmistakably to a lost intervocalic consonant and hence to an original **saCats*, which destroys the match. Naturally, our first guess would be **sanats*, but this time the comparative evidence provides no explicit support for an original nasal.

[255.] Bq **samin** 'bitter'

Bq **samin** 'bitter' : PNC ***ħimić'ć'^{w}V̌** 'sour' : Hattic **zzibina-** 'sour' (B9lb [58])

This proposal strikes me as more than a little hopeful, especially since Basque *m* cannot be etymological.

[256.] Bq **santan** 'ever': See under **senton** 'old man'.

[257.] Bq **sasi** 'bramble'

Bq **sasi** 'bramble' : PNC ***cā̲ce** 'burr' : Bur **čhAš** 'thorn' : PY ***seʔs** 'larch' (B9lb [27]) (B9lc [37]) (SR94 [148])

The word *sasi* is attested in the early seventeenth century as *çarci* (= *zarzi*); the development of *zarzi* to **zasi* would have been regular, as would the development of **zasi* to *sasi* by the usual Basque sibilant harmony. The Basque word can thus hardly be separated from Spanish *zarza* 'bramble', and the match is destroyed.

[258.] Bq **seme** 'son

Bq **seme** 'son' : PNC ***ćămV** 'kinsman, relative' : Bur **ǯaam** 'relation, kinsman (distant)' (B9la [136]) (B9ld) (B9ld PS1)

Basque *seme* should be from **senbe*, and *SEMBE* is in fact attested in Aquitanian. The word is possibly a compound of *sehi* ~ *sein* 'boy' (< **seni*).

[259.] Bq **senton** 'old man', **sentana** 'old woman', **santan** 'ever'

Bq **sen**-ton 'old man', **sen**-tana 'old woman', **san**-tan 'ever, never' : PNC ***śüno** ~ ***ś(w)äno** 'year' : Bur **šini** 'summer' : PY ***sin-** 'old' : Sum **sun** ~ **sumun** '(be) old' (B91a [128]) (B9lc PS1')

With their medial *-nt-* and their unorthodox sex-marking, the first two Basque words cannot possibly be ancient or native. They are confined to the French Basque Country, and are doubtless derived from some development of French *cent* ' 100' or its Occitan equivalent. Particularly interesting is the citing of French Basque *santan* 'ever, never' as representing an alleged root **sVn-* 'old; year'; this word occurs only in the French Basque Country in the phrase *sekulan santan* 'never in a million years' (*sekulan* 'never'), and it is a borrowing of French *cent ans* 'a hundred years', with the usual Basque sibilant harmony.

[260.] Bq **sirats** 'nerve, vein'

Bq **sirats** ~ **sirax** 'sinew, nerve, vein' (cf. dialectal i-**zorro** 'root') : Bur **cʰirIš** ~ **cʰiriiš** 'root' (cf. **-cʰiir** 'gut, entrail') (B9la [107]) (B9lc [140])

Basque *sirats* (and variants) is sparsely but securely attested in the sense 'nerve', 'vein', or 'tendon'. In the French Basque Country, an identical word means 'luck, chance, fate', and also 'intention, consolation, pleasure' and even 'magic spell, evil eye'. It is conceivable that these other senses derive from metaphorical uses of the anatomical term. The purely Roncalese word *izorro* 'root' cannot possibly be related to any of the other words.

[261.] Bq **sits** 'moth'

Bq **sits** 'moth' : PNC ***šʷɨ̱ʒV** 'bug, tick' : Bur **šišAr** 'a kind of beetle': Sum **ziz** 'moth' (B9lc PS2 [15])

Western Basque *sits* means both 'moth' and 'dung, ordure'; other varieties have *sats* in both senses. It is hard to know what to make of this, since there is no parallel elsewhere for an *a* ~ *i* alternation of this kind; we might posit an original **saits* or **seits,* but no such form is attested anywhere — nor is the semantic connection between 'moth' and 'dung' at all clear.

[262.] Bq **soin** 'shoulder, body, torso'

Bq **soin** ~ **soiñ** ~ **sün** 'body, upper body, bust', **soin**-ez 'bodily' : PNC ***čɨnHV** 'game, animal' : Bur **šAn** 'limbs, all the parts of the body' : Sum **su** ~ u-**zu** 'flesh, body' (B9la [43]) (B9lc [40])

The most conservative form of the Basque word is *soin*, and there is no point in citing either the secondary variants or the transparent derivative *soinez*. According to region, the word variously means 'surface of the body', 'body', 'upper body, torso', 'shoulder', 'load (carried on the shoulders)', or 'clothing', and in places the word denotes a specific article of clothing; of these, 'surface of the body', 'clothing', and 'shoulder' are the earliest attested senses.

[263.] Bq **su** 'fire'

Bq **su** 'fire' : Abkh a-m-**ca**, Abaza m-**ca**, Kabardian ma-**fă**, Adyghe ma-**śʷă**, Ubykh mə-**ǯe**, PND ***c'ajɨ** (C85 [85])

Bq **su** 'fire, hearth' : PNC ***c'aji** 'fire' : Bur **-ce** ~ **-ci** ~ **-Acii**-As 'set fire to, burn (tr.)', **ši** 'fireplace', **šu**-tin 'hearthstones', (W) **šu**-tum 'hearth, stone' : Sum **i-zi** 'fire, heat', **i-i-zi** ~ **na-i-zi** 'smoke' (B9la [65]) (B94b [35])

Basque *su* means only 'fire', never 'hearth'. The Basque-Caucasian resemblance has been pointed out several times before, for example by Uhlenbeck (1924:23) and by Bouda (1948:§10).

[264.] Bq **sudur** 'nose', **sunda** 'stench, rank smell', **sunbin** 'sob'

Bq **su**-dur(r) ~ **su**-gur(r) 'nose', **sun**-bin 'sob' (v.), **sun**-da 'smell, aroma' : PNEC ***s̄un**-t'V 'scent, odor; snuff; smell (v.)' : Bur **šuuŋ**(etAs) 'smell, blow (v.)' : PW ****ČUN(G)A** 'nose; smell (v.)' (B9la [21]) (BR94)

The words for 'nose' and 'smell' (the second actually means 'stench' 'rank smell', rather than 'smell' in general) have been arbitrarily segmented by the authors, and there is no reason within Basque to suppose that the two words are related, though of course they might be, at some considerable remove. The exceedingly obscure hapax *sunbin* is unlikely to be connected to the other words.

[265.] Bq **suge** 'snake'

Bq **suge** ~ **sube** 'snake' : Proto-Lezgian ***č'ek'-** 'fish, lizard' : PY ***cīk** 'fish, snake; meat' (B9la [93]) (B9ld PS1) (SR94 [241])

[266.] Bq **-t** (first-person singular agreement suffix in verbs)

Bq **-t** 'I' (ending of verbs) : PNC ***tV-** ~ ***tʰV-** ~ **-itʲi** 'we' (B9la [141]) (R94b)

The Basque agreement marker *-t* alternates with *-da-*, which occurs whenever any morphological material follows, and there is not the slightest doubt that **-da* is the original form of the morph. The proposed match is far-fetched. See the remarks under **ni** 'I' above.

[267.] Bq **-ta** ~ **-eta** (empty morph)

Bq **-ta** (locative) : PNC ***-tV** (lative) : Bur **-Aṭe** / **-eṭe** (adessive) (B9ld)

First of all, Basque *-(e)ta* is not "locative", and is not even a case-ending at all. It is an empty morph which precedes certain case-suffixes in certain circumstances, notably in certain plural forms: for example, *mendian* 'on the mountain' but *mendietan* 'in the mountains', both with locative *-n*. The origin of the morph is unknown, but it is noteworthy that an identical suffix *-eta* is exceedingly frequent in place-names, with the apparent sense of 'abundance (of)'. This observation has led Schuchardt and others to propose that *-eta* derives from the Latin morph -ETA, plural of -ETUM, which, if correct, would make the Basque affix cognate with Spanish *-eda*. In any case, there is not the slightest reason to suppose that *-(e)ta* was ever a case ending. See Michelena (1971b) for an essay on this morph.

[268.] Bq **toska** 'kaolin', Z **toxka** 'clod'

Bq **toska** 'clay' : PY ***təq-** 'clay/dirt' (SR94 [44])

With its initial *t,* Basque *toska* ~ *toxka* cannot possibly be a native word. It is, in fact, an obvious loan from the Romance word which appears in Castilian as *tosco* 'coarse, rough, crude' and throughout Iberian Romance and

Occitan in similar forms; Corominas and Pascual (1980) derive all of these, including the Basque word, from Latin TUSCU 'unrefined, dissolute'.

[269.] Bq **tu** 'spit, saliva'

Bq **thu-** ~ **to** 'spit' : PNC ***tujV** or ***tuk'** 'spit' : Bur **thū** 'spit' : PY ***duk** ~ ***duq** 'saliva' (B9lc [l48]) (R94b) (SR94 [223])

Basque *t(h)u* ~ *ttu* 'spit' (the hyphen is uncalled-for), with its initial *t*, cannot possibly be ancient, and it is clearly of imitative origin, as most likely are the other words.

[270.] Bq **txahal** 'calf'

Bq **txahal** [čahal] 'calf, heifer' : PNC ***ʔɬɨć'ʷilV** 'calf, heifer' (B91c [57])

Basque *tx* was not anciently a phoneme, but only an "expressive" variant of an ordinary sibilant, confined to diminutive and other expressive variants of lexical items; it is therefore pointless to try to find systematic correspondences for it. The comparative evidence for the word *txahal* shows clearly the ancient intervocalic nasal, and Michelena (1949a:487; 1977:303) therefore reconstructs ***xanal**, which destroys the match.

[271.] Bq **txar** 'bad'

Bq **txar** 'bad, weak' : (Cauc) Lezgian ***č'irɨ** 'wild' : Urartean **c'ir**-ab 'empty, uninhabited' (B9lb [57])

This proposed matchup, such as it is, depends crucially upon the presence of the initial *tx-* in the Basque word. But Basque *tx* is never etymological, and it isn't here. The word is merely a palatalized form of northern *tzar* 'bad', which in turn is no more than a specialized variant of *zahar* 'old', reduced to *-(t)zar* in postposed position (Michelena 1964b:131).

[272.] Bq **txiki** 'small'

Bq **txiki** 'small' : PAT ***č'ək'ʷə-** 'small' (C85 [40])

Bq **txiki** (dialectal) 'small' : PNC ***x̌ik'V** 'short' : Bur **čiki** 'small' : Sum **sig** 'small, thin', etc. (B9lc PS2 [49]) (B9ld)

There are several difficulties here. Basque *txiki* is again a palatalized form, this time from original *tiki*. This is a western form, eastern varieties having *txipi* or *ttipi* from *tipi*. It is impossible to be certain which of these two represents the earliest form of the word, but Michelena (1955:286) notes that the forms with *p* are overwhelmingly dominant in medieval documents (where the word often occurs as a surname or a sobriquet) and even in sixteenth-century texts. Hence the best guess is that *tipi* is earlier than *tiki*, which is not at all helpful to C. Moreover, words in [tʃi-] meaning 'small' or 'child' are omnipresent, not just in the Iberian Peninsula, but in the whole of western Romance, even though no Latin etymon can be identified: see, for example, Corominas and Pascual (1980) under *chico*. Indeed, words for 'small' in [ti-] or [tʃi-] are so common in the languages of the world, and so obviously motivated by phonesthetic factors, that such words can hardly serve as evidence for a genetic link. Consider English: Middle English *tine* 'very small', of unknown origin, developed via the Great Vowel Shift into modern *tiny*, as a result of which English speakers have altered the word into *teeny* to maintain the phonesthetic relation.

[273.] Bq **tximeleta** 'butterfly'

Bq **txime**-leta 'butterfly' : PNC ***ć'VmV**-ć' 'butterfly' : Bur **čumū**-so 'grub, fish insect' (B91c [58])

Basque *tximeleta* is just one of an extraordinarily large number of different regional words for 'butterfly', some of them transparent compounds and most of the others of apparent imitative origin. Today *tximeleta* is the most frequent word in the west, but the word has no early attestations: even Azkue, in his 1905 dictionary, does not record it, citing instead *mitxeleta* as the usual western form. It is out of the question that the unrecorded *tximeleta,* alone among the numerous words for 'butterfly', should represent the continuation of a word from six or seven thousand years ago: the word is nothing but a recent metathesis of *mitxeleta*, which in all probability is a derivative of *Mitxel* 'Michael' (Basque has very many names for insects, birds, fish, and other small living creatures which are derived from personal names).

[274.] Bq **txingurri** 'ant'

Bq **txingurri** 'ant' : PNC ***ʒimiʒV** 'ant' : Albanian **thënegulë** 'ant' (B9ld PS2)

The most conservative of the several variants of the Basque word for 'ant' is probably *iñ(h)urri*, found at both ends of the country. Central dialects have *txinaurri*, *xinhaurri*, *txindurri*, or *txingurri*. What is clear is that the initial *(t)x-* of some forms must be a late addition, one of the uncommon instances of palatalization by prefixation. This destroys the match with Caucasian, such as it was. Bengtson maintains that the Albanian word is a loan from "Macro-Caucasian".

[275.] Bq **txori** 'bird'

Bq **txori** ~ **xori** 'bird' : PNC ***č'wi̯lV** 'quail, bird, small bird' : Bur **čili** (babu:k) 'a very small bird' (B94b [33])

As always, the initial palatal cannot be etymological in Basque. Michelena (1955:275) concludes that *(t)xori* is merely the palatalized form of *zori*, which today means 'luck, fortune' but which is attested in the sixteenth century as 'omen'; he explains this development by appealing to the common practice of foretelling the future by observing the flight of birds. (Compare the Latin *auspex* 'augur', from **avi-spek-* 'observer of birds', and more particularly Old Spanish *auçe* 'luck, fortune' < Latin *AVICE 'bird'.)

[276.] Bq **txorru** 'root (of a hair)'

Bq **txorru** 'stubble (of hair or beard)' : PNC ***č'aHVrV** 'hair' : Bur **-čho:qur-** 'forelock' (B93)

Basque *txorru* does not exactly mean 'stubble'; it means 'root' or 'shaft' of a hair or whisker, and this only in B: elsewhere, the word means 'granary' or 'greenfinch' (this last may well be unrelated). As always, that *tx* cannot be etymological, and this is a diminutive of something, possibly of the universal *zorro* 'sack, bag, pod'.

[277.] Bq ****-tzi** 'ten'

Bq ***tzi** 'ten' : PAA *(b)ć'ʷV, PND ***u̯enc'i** 'ten' (C85 [21])
Bq ***-tzi** 'ten' : PNC *w-**enc'i** 'ten' : PY ***tuʔ**-n 'ten' (SR94 [271])

The alleged Basque **tzi* 'ten' is pure fantasy: no such word or stem exists or can be reconstructed in Basque, in which the universal word for 'ten' is *hamar*. Čirikba and Bengtson have extracted this morph from the numerals *bederatzi* 'nine' and *zortzi* 'eight' on the assumption that it must represent an ancient numeral for 'ten'. Now *bederatzi* is certainly an interesting word: it is much longer than the other numerals, and it shows every sign of being an ancient compound . Michelena (1977:134) is inclined to see the first element as **bade*, the form which he reconstructs for *bat* 'one', a conclusion accepted by C without discussion. This leaves an apparent morph *-(e)ratzi* unaccounted for. C and B have leapt to the conclusion that the whole word means 'one from ten': hence their **tzi*

for 'ten'. But this is sheer guesswork — neither *-*(e)ra* 'from' nor *-*tzi* 'ten' finds any support anywhere, and anyway the order is wrong: the rigorously head-final Basque expresses 'one from ten' as 'ten-from one'. (There is an eastern word *bedera* 'one apiece', but it is hard to see how this is helpful to C and B.) Nor is this speculation the only one on the table: others have noted that the morph *-eratzi* looks strikingly like a verb form, and suggested that the original meaning of *bederatzi* might have been something like 'one left', with a now lost verb.

This speculation is not original here: both Lafon (1933:167) and Bouda (1948:§115) offer the same interpretation.

The form *zortzi* 'eight' offers no real support: C and B are obliged to assume that **zor* must represent another lost numeral, this time meaning 'two'. Now, if we combine C's proposals here with his interpretation of *hamaika* (see above under ****-ika**), we find that he is making the following claims:

(1) Basque anciently had the numerals **ika* 'one', **zor* 'two', and **tzi* 'ten'. These were later replaced by the modern forms *bat* (< **bade*), *bi*, and *hamar*, respectively — even though the numerals for 'one' and 'two' are normally among the linguistic items most resistant to replacement.
(2) The numeral *zortzi* 'eight' was formed with the old numerals for both 'ten' and 'two' before they dropped out of the language.
(3) The numeral *bederatzi* 'nine' was formed with the *old* numeral for 'ten' but the *new* numeral for 'one'.
(4) However, the numeral *hamaika* 'eleven' was, in complete contrast, formed with the new numeral for 'ten' and the *old* numeral for 'one'!

I find it impossible to make any sense of this collection of implausible and conflicting conjectures: they cannot all be right, and I can see no reason to believe that any of them is right. The plain fact is that the Basque numerals offer no support for a conjectured link with Caucasian languages.

[278.] Bq **-tzu** (indefinite plural suffix), **-tsu** 'full of'

Bq **-(t)zu** (old plural suffix), **-tsu** (suffix of abundance) : Abkh **-c^{w}a** (plural suffix), PND ***-cwV** (C85 [5])

There is no justification for regarding the suffix *-(t)zu* as any older than the definite plural suffix *-k* (see under **-k** (plural) above). The point about this suffix is that it only forms *indefinite* plurals: *bat* 'one; a', *batzu(k)* 'some, several'; *nor* 'who?', B *nortzuk* 'who?' (plural); *hiru* 'three', *hirutzuk* 'group of three'. The fact that *-(t)zu* is so often reinforced by the more usual plural suffix *-k* might suggest that *-(t)zu* is a residue of an ancient plural suffix, but it certainly doesn't prove it. Note that this suffix is often described as "collective", rather than "plural".

The suffix *-tsu* has an alternative form *-zu*, which is almost certainly the original form; the variant *-tsu* seems to have developed from derivatives like *satsu* 'filthy' (< *sats* 'ordure' + *-zu*). This suffix forms adjectives from nouns, and the derivatives mean 'having an abundance of', like English *-ful* or *-ous*. It is impossible to tell whether the two Basque suffixes are related: in spite of an obvious semantic connection, they are very different in their behavior.

Exactly the same parallels are pointed out by Lafon (1944:529-530; 1951§9).

[279.] Bq **uda** 'summer'

Bq **uda** 'summer' : Sum **ud, utu** 'sun, day' (B9lc PS2 [33])

[280.] Bq **uhain** 'wave' (n.)

Bq u-**hin** 'wave' : PNC ***x̄ün?i** ~ ***x̱än?ɨ** 'water' (B9lb [20]) (B9lc [96], PS1') (B94b [19])

Basque *uhin* is merely a local variant of the more usual *uhain,* which is obviously a compound of *ur* 'water' (whose combining form is regularly *u-*); the second element is most likely *gain* 'top'. The connection with *ihintz* 'dew' proposed by Bengtson is fantastic; see under **ihintz** above.

[281.] Bq **ukabil** 'fist'

Bq u-**khab**-il 'fist' : PNC ***qʷa̲lpV** ~ ***Gwă̲bV** 'paw, wing' : Bur **qAf** 'claws' (B9lc [101]) (B93)

With Basque *ukabil,* we have an undeniable example of an erroneous arbitrary segmentation: the word is a transparent compound of the archaic *uko* 'forearm' (combining form *uka-*) and the element *-bil,* which expresses roundness. The first occurs in several other formations, such as *ukondo* 'elbow' (*ondo* 'below', 'next to'); the second is found in a sizeable number of formations, such as *barrabil* 'testicle' and *gurpil* 'cartwheel' *(gurdi* 'cart').

[282.] Bq **ukarai** 'wrist'

Bq u-**kharai** ~ u-**khare**, u-**khal-** 'wrist', u-**khai** 'forearm' : ECauc ***q'we̲lV** 'hand' (B91c PS1) (B93) (B94a [16], [22]) (B94b [26])

This item produces a similar missegmentation. Basque *ukarai* is another compound of the archaic *uko* 'forearm', this time with *garai* 'high part' (Michelena 1977:227). Bengtson cites the combining form *ukal-* because he finds it convenient in obtaining a match with Caucasian, but note that he does *not* cite the combining form *gal-* of **gari** 'wheat' (above), because there it is the *r* he wants to match. Once again, it is completely out of order for him to pick and choose among alternants in order to select the one he finds convenient.

[283.] Bq **ukondo** 'elbow'

Bq u-**k(h)ondo** 'elbow' : PND ***q'wV̲ntV** 'elbow, knee' : PY ***gid** 'elbow/joint' (B9lc PS1) (B93) (SR94 [79]) (R94b)

Here we have another spectacular misanalysis. Basque *ukondo* is an obvious compound involving *ondo* 'bottom, below, next to' (itself a loan from a Romance development of Latin FUNDU); the first element looks like the archaic *uko* 'forearm', but the existence of a variant form *ukalondo* leads Michelena (197lb) to conclude that this first element is actually *ukarai* 'wrist', itself a compound of *uko* with *garai* 'high place, above' (see the preceding item).

[284.] Bq **ur** 'water', **euri** 'rain'

Bq **ur** 'water' : PND ***ħwirə** 'river', 'lake' (C85 [83])

Bq **(h)ur** 'water', e-**uri** 'rain', **hur**-alde 'flood', **hur**-(xuri) 'urine', **hur**-(egin) 'sweat (v.)' : PNC ***Rʷer**-tV 'rain', ***ħwiri** 'lake' : Bur **hAr**-(aalt) 'rain', **hər**-Aš 'urine', **hur**-(ooγo) ~ **hur**-ogo 'sweat', **huur** ~ **hUr** 'conduit for water', **hUr**-(iginAs) 'stream, current, wave' : PY ***xur**(-es) 'rain' : Hattic **ura** 'spring (of water)' (B9la [85]) (B9lc PS1) (B93) (B94b [18]) (SR94 [212])

The Basque word for 'water' is *ur*, even in the aspirating dialects, save only in Z, the dialect which has extended the aspiration to all monosyllables which can bear it. As always, the *h* is not etymological, and there is no point in looking for cognate segments. There is no evidence that *euri* is related to *ur*, while the remaining Basque words are transparent compounds of *ur* and should not be cited separately. Bouda (1948:§102) also compares Basque **ur** with several North Caucasian words for 'river', 'spring', 'pond' and 'lake'.

[285.] Bq **urre** 'gold'

Bq **urhe** 'gold' : Berber **uriγ** 'yellow' or Laz **orXo**, Georgian **oXro** 'gold' (B94b)

Bengtson sees the Basque word as a loan from Berber or from Kartvelian. Naturally, many other scholars have tried to interpret it as deriving from Latin AURU 'gold', but the phonological difficulties are formidable.

[286.] Bq **urtxakur** 'otter'

Bq **urtxakur** 'otter' : PND ***tend̄ʷʌ** ~ ***teng̃ʷʌ** 'weasel, marten' : PY ***täχʌr** 'otter' (R94b) (SR94 [196])

For PY **täχʌr* 'otter', Ruhlen amazingly adduces Basque *urtxakur* 'otter'. But this purely Bizkaian word could hardly be more transparent: it is a compound of *ur* 'water' and *txakur* 'dog'. This compound can't even be of any great antiquity, since *ur* appears in its full form, instead of in its normal ancient combining form *u-*, while the word for 'dog' appears in its modern palatalized form instead of in its original form *zakur*.

[287.] Bq **uxuri** 'urine'

Bq **uxuri** 'urine' : Abkh **-čx̌ʷa-ra**, PND ***čV̌xwV** (C85 [54])

Basque *uxuri* 'urine (of an animal)' is a rather localized word; far more common in the sense of 'urine' are *pix(a)*, *piz(a)*, and *txiza*. Anyway, *uxuri* is a mere variant of the more usual *urxuri*, a compound of *ur* 'water' and *xuri*, the diminutive of *zuri* 'white' (see under *ur*, above), and it can be of no great antiquity.

[288.] Bq **uzki** 'anus, buttocks'

Bq u-**zki** 'anus', (bi)-**zka**-r(r) 'back' : Abkh a-**zkʷa** 'back' : Bur **-sqa** 'on one's back' (B9la [39]) (B9lc PS1) (B93)

Basque *uzki* means 'anus' in the west but 'buttocks' in the east. It is a transparent derivative of *utz* 'breaking of wind', formed with the common element *-ki*, which derives concrete nouns. And *utz* in turn appears to be a variant of *putz* 'puff of breath', 'fart', which is probably of imitative origin. The comparisons are fantastic, as is the introduction of Basque *bizkar*.

[289.] Bq **xahu** 'clean'

Bq **xanhu** 'clean' : PY ***tur-** ~ ***tul-** 'clean' (SR94 [45])

Ruhlen cites an alleged Basque **xanhu* 'clean', but the correct form is *xahu*, and, as the nasal vowels in the eastern forms show, this is nothing but a loan from Latin SANU (Michelena 1958:229) which has undergone the common "expressive" palatalization of Basque (cf. *xako* 'leather bag for wine' from Castilian *saco* 'sack, bag' or a similar Romance form).

[290.] Bq **-xe** (intensive suffix)

Bq **-xe** 'self' : PNC ***ǯʷV** 'self' : Bur **ǯi** 'life, soul, self' : PY ***ʔeʔte** 'alive': Sum **zi** ~ **ši** 'life. soul' (B91c PS2 [l])

Basque *-xe* does not mean 'self'. It is an intensive suffix: *orain* 'now', *oraintxe* 'right now'; *hau* 'this', *hauxe* 'this very one', *hemen* 'here', *hementxe* 'right here'. The same suffix is also widely attested in several other functions: excessive (*handi* 'big', *handixe* 'a little too big'), diminutive (*aita* 'father', *aitaxe* 'grandfather', lit., 'little father'), 'almost' (*aitu* 'finished' [participle], *aituxe* 'almost finished'). It is very difficult to determine the historical development of this odd morph, but note that one or two of its functions are consistent with a derivation from *xe* ~ *xehe* 'small' (< **zene*).

[291.] Bq **-z** (instrumental case-suffix)

Bq **-z** (instrumental suffix) : Abkh **-(a)s**, PND ***-s̄(ē)** (instrumental) (C85 [3])

Bq **-z** (instrumental) : PNC ***-s(ē)** (instrumental), Abkh **-(a)s** (instrumental), Chechen **-s** (animate ergative) : (Yen) Ket **-aś** (instrumental/comitative), Kott **-os** (comitative), **-s(e)** (instrumental) : Hurrian **-(u)s** (ergative) (B9ld) (B93) (B94a [32])

The same or similar resemblances are pointed out by Trombetti (1925:§§95-96), by Dumézil (1933:128) and by Lafon (1951:§3).

[292.] Bq **zahar** 'old'

Bq **zar** 'old' : PAA ***ž**^{w}, PND ***swɨ̄rHo** (C85 [36])

Bq **zahar** 'old' : PNC ***s**^{w}**ɨ̄rHo** / ***rɨ̄hɨ̄s**^{w}**o** 'old, year' : Bur **chōr** 'ancient, former(ly)' (B9lc PS1) (B9ld)

The form *zar* is the western contraction of the more conservative *zahar*. An intervocalic *h* is often an indication of a lost former consonant, but in this case no clear evidence of such a lost consonant can be found. Indeed, Michelena (1961) suggests that the form *SA.HAR*, which occurs in an ancient and puzzling inscription found in Lerga (Navarra), may represent this Basque word. None of this is of any assistance with the Caucasian matchup, which appears to require multiple metatheses.

Bouda (1949:§48) also compares the Basque word to North Caucasian forms.

[293.] Bq **zain** 'nerve, blood vessel'

Bq **zaiñ** 'vein, nerve, root' : PNC ***sē̲mhIV** 'sinew, muscle, intestine' : Sum **sa** 'sinew, rope' (B9la [55]) (B9lc [59])

[294.] Bq **zakar** 'strong, vigorous, brave'

Bq **zak(h)**-ar 'strong, violent, coarse' : PNEC ***c'aq'q'V** 'strength' : PY ***sak**-ar 'hard, stiff' : Sum **zag** 'power, might, strength' (B9la [113])

Basque *zakar is* both a noun meaning 'scab' and an adjective meaning 'clumsy, vulgar, coarse'. There is also a word *zangar ~ zankar,* which is a derivative of *zango ~ zanko* 'leg' (a Romance loan); as noun, it denotes various parts of the leg, while as an adjective it means 'strong, vigorous' and, by extension, 'brave' and also 'violent'. Lhande (1926) concludes that the adjective originally meant 'strong in the legs ', which seems eminently plausible. This word has a variant *zakar,* and that appears to be what Bengtson is citing; if so, the comparison can be dismissed out of hand.

[295.] Bq ****-zaki**

Bq **zaki** 'bone' (lepo-**zaki** 'neckbone, nape') : PDC ****c'aq-** 'bone' (B9la [58])

Basque *lepozaki* does not mean 'neckbone': it means 'back of the neck, nape', and also 'hair bun'. Azkue suggests 'bone' for *zaki*, but this no more than a wild guess, as his entry makes clear; I would suggest that it's not even a very plausible guess.

[296.] Bq **zakur** 'dog'

Bq **zakhur** 'dog', diminutive **txakhur** 'small dog, puppy' : (Cauc) Budukh **cäkul** 'fox' (?) : Bur **ǯAkun ~ ʐ̣Akun** 'donkey' (also Spanish **cachorro** 'puppy'; Sardinian **ǯáɣuru** 'hunting dog'; Greek dialect ζαγάρο-ς ~ ζαγάρι 'hunting dog') (B9ld PS2)

There is no doubt that Basque *txakur* is a diminutive of *zakur*, though today in most of the country *txakur* is the unmarked word for 'dog', while *zakur* is either absent or specialized to mean 'big dog'. And *txakur* absolutely never means 'puppy'. The idea that Spanish *cachorro* is derived from Basque *zakur* is an old one, but it rests upon no evidence, and it is dismissed as unsustainable by Corominas and Pascual (1980).

[297.] Bq **zamar** 'sheepskin, sheepskin jacket'

Bq **zam**-ar(r) 'lock of wool, shock of hair' : PNC *He-**c'ʷemV** 'eyebrow' : Bur **-ltaan**-c 'eyebrow' : PY ***cəŋe** 'hair' : PW ****TSUMA** 'hair' (B9la [l2]) (B9lc [43]) (BR94 [26])

Basque *zamar* does not mean 'lock of wool' or 'shock of hair'. It means 'sheepskin cloak', of the type traditionally worn by Basque shepherds as a raincoat, but *specifically one from which the wool has been removed*, and in some places it also means '(sheared) sheepskin' in general. The gloss 'lock of wool' is given in Aulestia (1989) but nowhere else; Aulestia's dictionary is popular and derivative, not scholarly; he provides no source or provenance for such a sense, and his report cannot be accepted in the absence of any evidence from scholarly dictionaries or from earlier written sources. This word is of unknown origin, but, once again, that medial *m* cannot be original. Finally, since the authors seem eager to get rid of that troublesome final *r*, I might point out that a derivation from Latin SABANU 'covering' cannot be ruled out (compare Basque *zamau* 'tablecloth', taken directly from the Latin word).

[298.] Bq **zango** 'leg'

Bq **zan**-kho ~ **zan**-go 'leg, foot' (**-kho** ~ **-go** diminutive) : PNC *He-**c'c'ʷemV** 'legbone, shin, ankle' : Bur **-ltAn**-c 'the whole leg', **-(l)ti(ī)n** 'bone' (B9la [56]) (B9lc [44])

The Basque word appears as *zango* in the central dialects but as *zanko* or *zankho* in both the west and the east; the word variously means 'leg', 'foot', 'calf' or 'paw'. The western forms with a voiceless plosive after *n* prove beyond dispute that this is a loan word from Romance; for a discussion of the origin of the word, see Corominas and Pascual (1980) under *zanca*.

[299.] Bq ****zaro**

Bq **zaro** 'night' (dialect) : PNC ***ś**ʷ**VrV** 'night' : Bur gōn-**šere** 'all night' (B9ld)

There is no such Basque word as **zaro* 'night'. Bengtson (or his immediate source) has apparently misread the entry in Azkue's dictionary, in which the editor is in fact explaining the confusion which had led some earlier scholars to misinterpret the derivational suffix *-zaro* as meaning 'night'.

[300.] Bq **zazpi** 'seven'

Bq **zazpi** 'seven' : Coptic **sašf** ~ **šašf** 'seven' (B94b)

Astoundingly, Bengtson sees the Basque word as a loan from ancient Egyptian. The word 'coincidence' appears not to be in his vocabulary.

[301.] Bq **zehe** ~ **xehe** ~ **ze** ~ **txe** 'small'

Bq **ze** 'small' : Bzedukh **-śa**, Abkh **-sa**, etc. (C85 [37])

Bq **txe** 'insignificant, small', **-txo/-txu** (diminutive suffix) : Abkhaz-Adyghe ma-**č'a** 'a little', 'small', Abkh -x^{w}-**č'ə** 'child, baby', x^{w}**č'ə**-k' 'a little' (C85 [39])

Čirikba is clearly unaware that Basque *xehe* ~ *txe* is nothing more than the palatalized form of *zehe* ~ *ze*, and he compares the two forms with different Caucasian words which they happen to resemble separately. Since the Z and R dialects exhibit nasal vowels in this word (Z *xẽhẽ*, R *xẽ*), it is clear that an intervocalic *n* has been lost from this word, and that the original form must have been **zene* (Michelena 1977:146), a form which does nothing to improve the match with Caucasian. Lexical items and affixes in Basque which have some semantic connection with smallness regularly undergo "expressive" palatalization; the presence of the segments *x* and *tx* is a consequence of this palatalization, and cannot be invoked to support a proposal of cognacy. The diminutive suffix *-txo* ~ *-txu* is not related to this item: it is merely the palatalized form of the ancient diminutive suffix *-to*, which is attested in Aquitanian.

[302.] Bq **zer** 'what?'

Bq **ze**-r 'what, how', **ze**-in ~ **zu**-ñ 'which', **ze**-n-bat 'how much', etc. : PNC ***s̄a** 'what', Bur be-**s** ~ be-**sA**-n 'what', be-**se** 'what' (and other interrogatives) : PY ***ʔas-** / ***sV-** (interrogative) 'what?' (B9la [146]) (B9lc PS1') (B9ld) (B94a [5]) (B94b [6]) (R94b) (SR94 [296])

Basque *zer* (which means only 'what?', and never 'how?') is exceptional in not being derived from the more widespread interrogative stem *no-*. The word *zein* 'which?' is merely a specialized development of *zeren*, the genitive of *zer*, while *zenbat* 'how much?, how many?' is in turn a derivative of *zein*.

[303.] Bq **zerren** 'destructive grub': See under **zorri** 'louse'.

[304.] Bq **zigar** 'mite'

Bq **zig**-ar(r) 'mite' (B9ld) : PNC ***ć^wĭkā** 'flea' : Bur **žiki** ~ **čeke** 'nit' (B9la [98])

[305.] Bq **ziho** 'tallow'

Bq **ziho** ~ **zigo** 'tallow' : PNC ***c'ēnxV** ~ ***c'ēnx^wV** 'fat' : Bur (W) **zAγ** 'cooked fat' (B9lb [15]) (B9lc [98])

The comparative evidence here yet again points indisputably to an original nasal, and hence to **zino* (Michelena 1950a:456).

[306.] Bq **zikiro** 'castrated goat'

Bq **zikhiro** '(castrated) goat' : PNC ***c'ɮ̲kV** 'goat, kid' : Bur **chiigir** 'nanny goat' : Sum **sigga** 'billygoat' (B9lc PS2 [5]) (B9ld)

[307.] Bq **zilbor** 'navel'

Bq **zil** 'navel, penis', **zil**-kho, **zil**-bor 'navel' : (Cauc) Tsakhur **ʒil** 'navel' (B9lb [11])

Basque *zil* has a simply enormous number of different meanings, of which 'navel' and 'umbilical cord' are just two. The obscure word *zilko* is a simple derivative of *zil*, while *zilbor*, a common word for 'navel' in the west today, is cited by Azkue (1905) only in the sense of 'abdomen'. There seems little reason to doubt the antiquity of the Basque word, but a resemblance to a word in a single Caucasian language can hardly count for anything.

[308.] Bq **zilar** 'silver'

Bq **zilh**-ar 'silver' : Berber a-**zerf** 'silver' (B94b)

Bengtson sees the Basque word as a loan from Berber. In fact, the vast majority of vasconists have long accepted that the Basque word derives from the same source as the Germanic word which appears in English as *silver*, though this etymology is categorically rejected by Agud and Tovar (1988-).

[309.] Bq **zizari** 'worm'

Bq **zizari** ~ **zizare** 'worm', diminutive **xixari** ~ **txitxari** : PNC ***xʷitxʷIɬV** 'lizard', ***čVrčVlV** 'lizard. snake' (B9la [94])

There is no point in citing the Basque diminutive separately, since it is only a regular derivative of the basic form of the word. It appears that Bengtson is invoking a clearly secondary form merely in order to improve the rather feeble match with Caucasian.

[310.] Bq **zorri** 'louse', **zerren** 'destructive grub'

Bq **zorri** 'louse', **zerr**-en 'moth, grub, maggot' : PNC ***sVr**-sV 'nit; germ (of an egg)' : Sum **šar**-in ~ **šur**-in 'a kind of insect, worm' (B9la [100])

There is no earthly reason to suppose that the two Basque words are related to each other.

[311.] Bq **zorrotz** 'sharp'

Bq **zar**-tzu, **ziri**, **zorrotz**, all 'sharp' : Abaza **c'ara**, Bzhedukh **-ć'ar**, Abkh **-c'ar**, all 'sharp' (C85 [35])

Once again, C is citing nonexistent words. I have never encountered either the alleged Basque word **zartzu* or the alleged **ziri*: the first does not exist at all, while the second exists in a variety of meanings none of which remotely approximates to 'sharp'. The universal Basque word for 'sharp' is *zorrotz*. And not only has C seemingly invented a couple of words, he has even confidently segmented one of them in order to get rid of the portion he can't match in Abkhaz. But what is this *-tzu*? The only suffixes approximately of this form are the indefinite plural suffix *-(t)zu* discussed under **-tzu** above, a rare and obscure suffix *-tzu* denoting 'profession', and the rather common adjective-forming suffix *-zu* ~ *-tsu* 'full of', resembling English '*-ous*' and also discussed above. This last at least forms adjectives, but its semantic relevance is far from clear, and anyway no relevant root **zar* is attested either.

Bouda (1952:§35) also compares Basque *zorrotz* with items in several North Caucasian languages, including Abkhaz.

[312.] Bq **zu** 'you'

Bq **zu** 'you' : Bzyb Abkh **śʷa**-ra, Abkh **šʷa**-ra, Ubykh **śʷə**-Rʷa, Adyghe **śʷa**, PND ***źwV** 'you' (C85 [13])

Bq **zu**, **zu-** 'you' (plural) : PNC ***źʷV** 'you' (plural) : Sum **za** ~ **zé** ~ **ze** 'thou', **-zu** 'thy' (B9la [144]) (B94a [4]) (B94b [5])

Oddly, Bengtson describes Basque *zu* as "plural", even though it is strictly singular in the modern language. Nevertheless, there is no doubt that this pronoun was anciently a second-person plural form. This resemblance is noted by Uhlenbeck (1924:578), by Trombetti (1925:87), by Lafon (1944:529), and by Lafon (1952a:§25).

[313.] Bq **zuhain** 'hay, fodder','tree'

Bq zu-**hain** ~ zu-**hañ** 'tree' : PNC ***k^{w}anʔV*** 'splinter, soft wood' and/or ***k^{w}Vn-tV̄*** 'crown of a tree' : Bur **kuna** 'rod, pole, stick' : Sum šu-**kin** (B9lc PS2 [6])

Though Basque *zuhain* is attested as 'tree' in Z, it more commonly means 'fodder, hay' or 'sprout, shoot' or sometimes just 'plant'. In any case, the comparative evidence allows Michelena (1977:304) to reconstruct the word as **zunai*, which destroys the match, such as it was.

[314.] Bq **zuku** 'soup, broth'

Bq **zuku** 'soup' (?) : PY ***ʔuʔχ*** ~ ***xuʔk*** 'soup' (SR94 [249])

In different parts of the country, Basque *zuku* means both 'soup' and 'broth' (the difference is culturally important in the Basque Country), and it also means, in places, both 'fruit juice' and 'slurry fed to pigs'. The word is an obvious loan from Latin SŪCU ~ SUCCU 'juice'.

[315.] Bq **zur** 'wood'

Bq **zur** 'wood' : PNC ***c'^{w}ɬhV*** 'tree, stick' : Bur **šAr** 'branch, bough' : Hurrian **sar**-me 'wood': Urartean **c'ara** 'orchard' (B9lb [24])

[316.] Bq **zuri** 'white'

Bq **zuri** ~ **xuri** 'white' : PNC ***h^{w}ö-č'^{w}örV*** 'gray, yellow' : PY ***suɾ-*** 'yellow' (B9la [124]) (SR94 [312])

The etymology of Basque *zuri* (of which *xuri* is a diminutive) is not known, though Azkue (1905) makes the interesting suggestion that it might be derived from *zur* 'wood' by means of the ancient adjective-forming suffix *-i*.

[317.] Bq **zuzen** 'straight'

Bq **zuzen** 'right (rectus), right (dexter), correct, straight, just', etc. (cf. **zin** 'oath, true') : Bur **chAn** 'straight, direct, true, right, correct' : PY ***tat(aŋ)*** 'straight' (B9lb [54]) (SR94 [258])

The primary meaning of Basque *zuzen* is 'straight'; it also has the transferred meanings 'correct', 'true', and 'just', but never 'right (hand)'. There is no reason to connect this word with *zin* 'oath'.

4. Evaluation and conclusions

This completes the survey of published data involving Basque and Dene-Caucasian which I have been able to find and consult. Before evaluating it, I need to consider some general points — above all, the authors' segmentation of Basque words.

In the great majority of comparisons, the authors have segmented the Basque words freely, removing any parts which they do not find convenient and consigning them to the outer darkness, in most cases without the slightest attempt at justifying this practice. (They do the same with words from the other languages, of course, but I won't pursue that here.) This happens most often with the very frequent initial vowels of Basque, which the investigators almost invariably find themselves unable to match, but in no case do these workers hesitate to remove any number of "prefixes" and "suffixes", even though all these morphs are unidentified and completely mysterious, and even though ancient prefixes are virtually unknown in Basque. In the case of compound words, it is often a simple matter to demonstrate that these arbitrary segmentations are preposterously wrong (look, for example, at **berezi** 'separate', **emakume** 'woman', **eskubarne** 'palm', **gurpil** 'wheel', **intzigar** 'frost', **lorratz** 'track',

makutsik 'in one's shirtsleeves', **ukabil** 'fist', and **ukondo** 'elbow', to cite just a few of the more egregious cases). With monolexemic words, this is not generally possible: I can't prove that **egun** 'day' was not anciently **e-* + **-gun,* or that **izter** 'thigh' was not **izt-* + **-er,* or that **lagun** 'companion' was not **lag-* + **-un.* But that's not my problem: the onus is clearly on the authors to defend these arbitrary segmentations. They attempt to do this only in certain cases, notably with their "prefixes". Čirikba and Bengtson, at least, claim the right to remove a "body-part prefix'' **b(e)-*. Let us consider that.

Now the idea that *b(e)-* is a prefix in names of body parts is indeed an old one, going back at least to Uhlenbeck (1927); it has been endorsed by several vasconists, including Michelena, who suggested that the prefix might represent a fossilization of the possessive *bere* 'his/her own'. However, this hypothesis rests on nothing more than the observation that the names of several body parts begin with *b(e)-*: *buru* 'head', *begi* 'eye', *belarri* 'ear', *beso* 'arm', *bihotz* 'heart', *birika* 'lung', *bizar* 'beard', *mihi* 'tongue' (< **bini*), *bare* 'spleen', *behazun* 'bile', *barrabil* 'testicle', *bular* 'breast', *belaun* 'knee', and *bizkar* 'back'. (I omit several obvious derivatives of *begi* 'eye'; I omit *behatz* 'toe, finger', because it appears to be a transparent compound of *behe* 'lower, below' and *hatz* 'finger'; and I omit *bekoki* 'forehead' because this is either a derivative of the loan word *beko* 'forehead' or a derivative of *begi* 'eye'.) That's it. That's all there is. On the other hand, there is no *b-* in the names of the majority of body parts: *ile* 'hair', *sudur* 'nose', *aho* 'mouth', *ezpain* 'lip', *hortz* 'incisor, tooth', *hagin* 'molar tooth', *lepo* 'neck', *eztarri* 'throat', *zintzur* 'throat', *ugatz* '(female) breast', *sorbalda* 'shoulder', *ukondo* 'elbow', *esku* 'hand', *hatz* 'finger', *eri* 'finger', *gerri* 'waist', *sabel* 'stomach', *hertze* 'intestine', *ipurdi* 'buttocks', *izter* 'thigh', *zango* 'leg', *oin* 'foot', *gibel* 'liver', *giltzurrin* 'kidney', *azal* 'skin', *soin* 'body', and many others.

All that we have, then, is the observation that the names of about fourteen body parts, among many others, begin with *b-*. Is this of any significance? In my view, certainly not. Basque has only a small inventory of consonant phonemes, and only a subset of these can occur in initial position in words which go back to pre-Basque: *b-*, *g-*, *l-*, *z-*, *s-*, and *n-*, and finally also *m-* in cases in which this derives from **b-*. And initial *b-* is *very* common. Azkue's 1905 dictionary, which omits recent loans from Romance, lists 74 pages of words beginning with *b-*, more than for any other consonant except *z-*, and this figure makes no allowance for the frequent development of original **b-* into *m-*; the once non-existent *m-* gets another 58 pages in Azkue's dictionary, suggesting that *b-* was anciently the most frequent word-initial consonant.

It is easy to find other areas of the vocabulary in which *b-* is more than averagely frequent. To take one example, four of the first ten numerals have names beginning with *b-*. To take another, the grammatical words and morphs of Basque include an astonishing number of items in *b-*: *bai* 'yes', *ba* (affirmative), *ba-* and *baldin* 'if', *(b)ere* 'also', *beraz* 'and so, consequently', *ber-* 'again, re-' (virtually the only native prefix involved in word-formation, and probably a calque), *be-* (jussive), *baino* 'than', *baina(n)* 'but', *baizik* 'but' (in 'not X but Y'), *bezain* 'as...as', *bezala* 'like, as', *bait-* (a verbal prefix which forms certain types of subordinate clauses), *-bait* (a suffix which forms indefinite pronouns), B *barik* 'without', northern *baita-* (a morph used in inflecting animate NPs), and the very common reflexive element *ber-* 'self, same' (I omit the indefinite article *bat* and the modal *behar [izan]* 'must', which are derived from a numeral and a noun, respectively). Of course, some of these are certainly or probably related to others, but the example nonetheless illustrates the great frequency of *b-*.

Furthermore, an ancient word-forming prefix **b(e)-* would be a striking anomaly in Basque: not a single word-forming prefix can be reconstructed with certainty for pre-Basque (though see Trask [forthcoming *a*] for an argument that there was a prefix for forming verbal nouns). Basque is, apart from a few prefixes involved in constructing verb forms, and two or three prefixes acquired recently under Romance influence, exclusively suffixing.

Finally, consider English. English has more than twice as many word-initial consonants as Basque, and no one has ever suggested a body-part prefix *b-* for English. Nonetheless, English has the body-part names *brain*, *brow*, *beard*, *breast*, *bosom*, *bust*, *back*, *belly*, *bowel*, *buttocks*, *bile*, *bladder*, *blood*, *bone*, and, of course, *body*. In short, English has just as many body-part names in *b-* as Basque.

I conclude, therefore, that there is no real evidence for an ancient body-part prefix **b(e)-* in Basque: the modest number of body-part names in *b-* is not a fact which requires an explanation. Readers are free not to share my skepticism, but I cannot see that anyone has made a good case for such a prefix. Hence the authors' frequent removal of this "body-part prefix'' in seeking comparisons should be viewed with deep suspicion, particularly since they do not hesitate to remove any other bits of body-part names which they find inconvenient (look at **behazun**

'bile', **belaun** 'knee', **bizkar** 'back', and **bular** 'breast'), because they happily remove the "prefix" from such implausible items as **bizi** 'alive', and above all because they do NOT remove their "prefix" whenever they find it more convenient to retain the initial *b-* (look at **bizar** 'beard', **beso** 'arm', **begi** 'eye', **bular** 'breast', and **buru** 'head'). In other words, *b-* is a prefix whenever they can't match it, but part of the stem when they need it to make a comparison go through. I find myself unable to take this seriously.

Apparently aware of the embarrassment of these arbitrary removals, Bengtson in some of his articles makes an ingenious attempt to deal at least with the initial vowels of Basque: he proposes that they are fossilized noun-class prefixes comparable to those found in many Caucasian languages (see especially Bengtson 1991c, 1993). Now this is certainly a clever idea, and, if Bengtson could make it work, he would have achieved an impressive breakthrough. Unfortunately, he can't. Apart from his "body-part prefix" **b(e)-*, which he expands as "body parts, fluids, and attributes", Bengtson proposes three such noun-class prefixes: **a-*, **e- ~ *i-*, and **u- ~ *o-*.

For his **a-*, he provides the following noun-classes: "persons, animals, body-parts, intangible nouns, trees, and artifacts". But this amazing list leaves almost no semantic stone unturned (how many nouns can you think of that wouldn't be covered by this list?), and he is clearly admitting that Basque has a lot of nouns in *a-* in all areas of the vocabulary, which we already knew. For his **e- ~ *i-*, he suggests "natural phenomena" in one paper but "things that are part of a larger collective" and "things that are (or appear) small" in the other. And for his **u- ~ *o-*, he proposes "body parts and fluids".

Each of these alleged prefixes he supports with a tiny number of examples: nine for **be-*, sixteen for **a-*, twelve for **e- ~ *i-*, and just six for **u- ~ *o-*. Even so, some of the examples he cites are certain or probable loan words: for example, **akain** 'tick', the wrongly cited ***ador** 'wheel' (correct **adorra** < **arroda**), **astigar** 'maple', **agure** 'old man', **bekoki** 'forehead'. Others do not seem to fit their alleged categories very well: 'milk' and 'drop' as "natural phenomena"? All but one of his examples of **u-* are transparent compounds of *uko* 'forearm': they all start with *uk-*, they all denote parts of the forearm, and they all have simple etymologies.

Three of Bengtson's four classes remarkably include "body parts" as one of their functions, the exception being **e- ~ *i-*. But what do we then say about *ile* 'hair', *ezpain* 'lip', *eztarri* 'throat', *esku* 'hand', *eri* 'finger', *ipurdi* 'buttocks', *izter* 'thigh', *hezur* 'bone', and *hertze* 'intestine'?

Moreover, it is a trivial exercise to adduce large numbers of nouns with any given initial vowel which do not remotely fit Bengtson's categories. For example, he wants "body parts and fluids" for his **u- ~ *o-*, but what about *otso* 'wolf', *ogi* 'bread', *ohe* 'bed', *ohar* 'observation', *oihan* 'forest', *ohol* 'wooden plank', *oihu* 'cry', *oker* 'error, injury', *okil* 'woodpecker', *ola* 'foundry', *olo* 'oats', *ondar* 'sand, beach', *orga* 'wagon', *orri* 'leaf', *ote* 'gorse', *uda* 'summer', *uger* 'rust', *uko* 'refusal', *ume* 'child', *untzi* 'vessel', *urde* 'hog', *urre* 'gold', *urte* 'year', *uso* 'pigeon', *usain* 'odor', *uzta* 'harvest, and many, many others?

Bengtson's idea is an interesting one, but he has provided zero evidence to support it. If he wants to make a case for fossilized noun-prefixes, then, at the very least, he needs to undertake a statistical study of selected areas of the Basque vocabulary, excluding obvious loan words and compounds, to see if any correlations appear. I am not optimistic about his prospects in this enterprise, but nothing less will do: merely citing a handful of confirming instances, while ignoring a much larger body of counterexamples, achieves nothing. (And what does he do about adjectives? Basque has lots of adjectives with initial vowels, too.)

I turn now to the proposed grammatical matchups. Shared grammatical morphemes, of course, constitute some of the most powerful evidence for genetic relations, and the authors propose more than two dozen shared grammatical morphemes involving Basque and at least one of the other languages.

The most determined attempt in this direction is found in Bengtson (1991d:164), who compares case-endings in Basque, North Caucasian, and Burushaski. Here is his tableau; I have silently replaced some of his labels for Basque cases with more accurate ones:

Basque	North Caucasian	Burushaski
-i [Dative]	***-i/*j** [Dative]	**-e** [Genitive]
-ra [Allative]	***-rV** [Locative]	**-ar** [Allative]
-ta	***-tV** [Lative]	**-Ate/-ete** [Adessive]
-n [Locative]	***-nV** [Genitive]	**-Ane/-A, e/-e e** [Comitative]

Basque	North Caucasian	Burushaski
-z [Instrumental]	***-s(ē)** [Instrumental]	
-k [Ergative]	***-k'V** [Instrumental,	**-Ak/-ek** [Instrumental]
-ik [Partitive]	Ablative, Partitive]	

First, as pointed out under **-ra** (allative) above, there is clear evidence within Basque that allative *-ra* is a secondary form derived analogically from the other allative form *-a*, which destroys the matchup. And, as pointed out under **-ta** (empty morph) above, Basque *-ta* is not a case ending at all, let alone a "locative" (as Bengtson would have it), and it is moreover probably of Latin origin. As for the rest, all I can see is a demonstration that languages with rich case systems often have some case-ending or other containing an unmarked segment like *-n*, or *-i*, or *-s*, or *-k*. The matchups offered are no better than arbitrary. Compare the rather modest case system of Turkish, where definite accusative *-i*, dative *-e*, genitive *-in*, and locative *-de* seem to match both Basque and Caucasian rather well (Turkish doesn't have an ergative, an instrumental, or a partitive). Even the case-marking postpositions of Japanese match up with all these languages surprisingly well: nominative *ga*, dative *ni*, comitative *to*, ablative *kara*, instrumental *de*, genitive *no*, and allative *e*. Indeed, it is not an easy task to find case systems which *don't* exhibit striking similarities to that of Basque: for example, the Kartvelian language Laz and the Muskogean language Koasati both share with Basque absolutive *zero* and ergative *-k*.

As for the other grammatical parallels offered, most are inconsequential or unacceptable. Basque **-a** (article) and **-antz** 'toward' are recent formations within the language; **d-** (verbal prefix) does not have the function imputed to it; the treatment of the three demonstratives (see under **hau** 'this' and **hau[r]** 'this') is an unholy mess; the comparison of **hi** 'you' depends upon an unsustainable correspondence; the comparisons involving **gu** 'we', **-k** (plural), **-la** (adverb of manner), **-n** (past tense), and **-ra** (causative) are too feeble to take seriously, and some of them involve comparisons with nothing but Abkhaz; the remarkable alternations involving **-k** (second-person agreement marker) are not paralleled in the other languages; the comparison of **-t** (first-person agreement marker) is phonologically and semantically indefensible; the comparison of the wrongly glossed **-xe** (intensive) is semantically hopeless.

Basque **ni** 'I' is matched only with Caucasian, even though first-person singular pronouns with nasals are enormously common in the world's languages. Similarly, **zu** 'you' and **-tzu** (indefinite plural) are matched only with a subset of Caucasian languages, and not with Caucasian as a whole (the Sumerian data adduced for **zu** do not in fact match). The interrogatives in **no-** and **ze-** are slightly more interesting, but many of the data adduced from other languages appear to require a leap of faith that we are actually looking at interrogative stems in *s-* and *n-*, rather than at fossilized case-suffixes. Once all the objectionable comparisons are removed, then, all we have is a few Basque grammatical morphemes that look vaguely like some grammatical morphemes in some other languages, usually only in Caucasian and sometimes only in Abkhaz. Once again, I cannot see that these vague resemblances constitute evidence for anything.

It remains only to consider whether the proposed matchups succeed in revealing any systematic correspondences. Recall that one of Michelena's chief reasons for dismissing the earlier work on Basque and Caucasian was the absence of even an attempt at identifying correspondences: segments were simply paired off with vaguely similar segments in other languages, at the whim of the analyst. Given the massive shortcomings already identified, the work under discussion here can only be regarded as an advance on that earlier work if it has indeed succeeded in uncovering some systematic correspondences .

Both Čirikba and Bengtson devote some attention to this issue in their various publications, but Čirikba's brief passage at the end of his paper succeeds only in demonstrating the absence of any such correspondences. Bengtson devotes more space to the matter, and one of his papers (199lc) in fact purports to be a demonstration of systematic correspondences among his "Macro-Caucasian" languages. Let us consider how successful he has been, always bearing in mind that in very many instances he has simply consigned to phonological oblivion any segments which he cannot match successfully. Since I lack the space here to provide an exhaustive examination of Bengtson's correspondences (which can be recovered by any interested reader from the data above), I propose to concentrate upon the sibilants.

In his table on pp. 147-148, Bengtson lists the following correspondence sets for the words considered in this particular paper; the numbers in square brackets are the numbers of the examples in this paper:

Basque	Caucasian	Burushaski	
s / -tz	**c**	**ch** (**čh** before **A**)	[34-39]
s	**č**	**š**	[40-42]
s	**š**	**š**	[63-66]
(t)s	**ć**	?	[54-56]
z	**s**	**s**	[59-62]
z / -tz	**c'**	**š**	[43-48]
z / -tz	**č'**	**š**	[49-53]
tx	**ć'**	**č**	[57-58]

Observe that most of these sets are exceedingly small, and they are in fact made smaller by the fact that many of the comparisons adduced involve only two of the three languages. Still, a fairly small number of very consistent correspondences is usually considered to constitute strong evidence in comparative linguistics. Is that what we have here?

Unfortunately, no. In set [34-39], Basque **s** is supposed to correspond to Cauc ***c** — yet the very first example (**sagu** 'mouse') matches Basque s with Cauc ***cʷ**. And the Burushaski segment is supposed to be **ch** or **čh**, but in two cases the segment cited is **c** (only one of these involves Basque: **bost** 'five'). Moreover, with Basque **sasi** 'bramble', the first **s** is indeed matched with Burushaski **čh**, but the second is paired with Burushaski **š**.

In the tiny set [40-42], Basque **s** in **itsaso** 'sea' is in fact matched with Caucasian and Burushaski *zero*. In set [63-66], which has no three-language comparisons, the initial **s** of **sahats** 'willow' is matched with Burushaski **š**, but the final **ts** is instead matched with Burushaski **s** — with no explanation. (Note that Bengtson does not in general distinguish a fricative from its corresponding affricate in setting up correspondences, which is in fact quite correct for initial and final positions in Basque.)

In set [54-56], the Caucasian segment is given as ***ć**, yet all three examples cited show ***ćʷ**. No doubt this is a typo, but it's unhelpful. In set [59-62], one of the three Caucasian forms shows ***sʷ** instead of the expected ***s**. Similar problems affect the rest of the sets: the segments are sometimes not quite what Bengtson says they ought to be.

No doubt a few of these are minor errors which could be fixed, but rather than continuing with the very small number of examples adduced by Bengtson to support his announced correspondences, I would prefer to examine the correspondences in *all* the data published by Bengtson. Again, I only have space here for a small selection. Let's consider the matchups proposed by Bengtson for the Basque sibilants in word-initial position in his various publications.

Basque **z-** is matched with the following segments; Caucasian segments in brackets are those adduced from some level below PNC:

Basque	Caucasian	Burushaski	
z-	***sʷ-**	**ch-**	**zahar**
	***s-**	—	**zain**
	[*c'-]	—	**zakar**
	[*c-]	**ǯ-, ʐ̌-**	**zakur**
	***c'ʷ-**	**-lt-**	**zamar**
	***c'c'ʷ-**	**-lt-**	**zango**
	***ɕʷ-**	**š-**	**zaro**
	***s-**	**-s**	**zer**

Basque	Caucasian	Burushaski	
	*c'w-	ž- ~ č-	**zigar**
	*c'-	z-	**ziho**
	*c'-	ch-	**zikiro**
	*xw-, *č-	—	**zizari**
	*s -	—	**zorri**
	*ź w-	—	**zu**
	*ć w-	š-	**zur**
	*č w-	—	**zuri**
	—	Ø	**zuzen**

Now this does not look to me remotely like a set of systematic correspondences: there are no two Basque words showing the same sibilant correspondences in Caucasian and Burushaski, and hardly any Basque words showing the same correspondences in Caucasian. But, of course, the small number of consonants in Basque and the huge number in Caucasian was bound to cause problems of this sort, as I pointed out near the beginning of this paper. Let us see what happens with Basque initial **s-**; this time the Burushaski correspondence in brackets is one which Bengtson is unsure is Burushaski at all:

Basque	Caucasian	Burushaski	
s-	*ǯ-	-s-	**sabel**
	[*č-]	š-	**sagar**
	*c w-	[čh-]	**sagu**
	—	š-	**sahats**
	*h-	—	**samin**
	*ś-	š-	**santan**
	*c-	čh-	**sasi**
	*ć-	ǯ-	**seme**
	—	ch-	**sirats**
	*š w-	š-	**sits**
	*č-	š-	**soin**
	*c'-	c- ~ š-	**su**
	[*s-]	š-	**sudur**
	[*č'-]	—	**suge**

This is no better — in fact, it's worse, since many of the Caucasian and Burushaski segments matched with Basque **s-** are the same ones which are also matched above with Basque **z-**. I can see nothing here that remotely deserves to be called a systematic correspondence. All we ever get is random similarities, lexical items which have been selected merely because they contain some kind of sibilant or other. The rule is clear: any sibilant can be matched with any other sibilant, according to the requirements of the thesis which the author is trying to demonstrate. And I cannot take this seriously: we are clearly looking at nothing but an assembly of chance resemblances between languages which have apparently been selected on some *a priori* basis as grist for comparison. I will not pursue this issue further here, but a patient reader can establish for himself that the matchups between sibilants in other positions, or between other segments, are no better.

It is time to summarize my findings. Of the 300 or so Basque items which have been adduced as "cognates" for words in North Caucasian, Burushaski, Yeniseian, Sumerian, or other "Dene-Caucasian" languages, more than half can be dismissed out of hand: they are obvious loans from Latin or Romance, or words which could not possibly have been in the language as recently as 2000 years ago, or words whose phonological forms in Roman

times would have been so different as to destroy the proposed matchups, or else they are severely localized regional forms, with no early attestations, to which no great antiquity can reasonably be assigned. In a few cases, the alleged Basque words do not exist at all. Moreover, many more Basque words have been wrongly cited, or wrongly glossed, or they have been preposterously mis-segmented. On top of this, almost all of them have been arbitrarily chopped up into pieces at the whim of the authors, in order to extract the bits that are considered convenient, while the remaining pieces, sometimes amounting to the larger part of a word, are unceremoniously dismissed from consideration as ancient "prefixes" or "suffixes" — even though prefixes are virtually unknown in Basque. In particular, the very frequent initial vowels of Basque are almost without exception removed from consideration, a policy which I hope I have shown is utterly without justification.

The attempts at comparing grammatical morphemes are no more successful: some of the Basque items adduced are of no great antiquity, and others are so short and formless as to constitute evidence for nothing. In no single case is any alternation in Basque paralleled by a similar alternation in any other language. All that we have is the banal observation that Basque and the other languages have some affixes or other involving unmarked segments like *-k*, *-a*, and *-n*.

Readers may decide for themselves whether the remaining items, involving Basque words of some real antiquity, constitute evidence for anything. For my part, I am certain that they do not: all I can see is a modest list of Basque words which bear a vague resemblance to some words in some other languages, of the sort that one can always find if one has sufficient patience and determination and a large dictionary. There are no systematic correspondences, but only random resemblances. The evidence for relating Basque to North Caucasian or to Burushaski is in no way more impressive than the joke demonstration of a genetic link between Basque and Hungarian found in Trask (1994). This demonstration contains more than 60 impressive "cognates" linking Basque and Hungarian, two languages which all the linguists whose work I have scrutinized here agree are not related at anything below the global level — and yet it took me only four hours to assemble.

There is an old adage in historical linguistics : "Look for Latin etymologies on the Tiber". The work under discussion illustrates the force of that adage as well as anything could. It is utterly pointless to extract Basque words from dictionaries and to compare them breathlessly with words from languages spoken in the Caucasus, in the Himalayas, in Siberia, and even in China and North America, without first scrutinizing the languages closer to home: in this case, Latin and the Romance languages, the neighboring languages which have indisputably had such a profound impact upon the vocabulary of Basque. Even in cases like Basque *moto* 'headscarf', for which I have been unable to identify a Romance source with certainty, which is more likely: that this phonologically very un-Basque word is a loan from an unidentified Romance source, or that it continues an ancestral Dene-Caucasian word from many thousands of years ago?

Finally, I remind the reader that this examination has focused exclusively on the Basque data: I have taken the data presented from other languages at face value. However, given the huge number of errors found in the Basque data, I am naturally inclined to wonder just how reliable those other data are. Given the authors' track record as regards Basque, is it likely that a Burushaski specialist, if any exist, would find the Burushaski data to be error-free? Or is it more likely that scrutiny of these data would also reveal a huge number of errors, thereby destroying yet more matchups?

As far as I can see, the evidence so painstakingly assembled for relating Basque to the other "Dene-Caucasian" languages amounts to precisely zero. Basque remains as isolated as it ever was.

Acknowledgements

I am indebted to Alexis Manaster-Ramer, Paul Sidwell, Marc Picard, Arantza Diez de Tuestra, José Ignacio Hualde, and Joseba Lakarra for variously bringing relevant work to my attention or for assisting me in obtaining it. None of these people necessarily shares any of the views expressed in this paper, and all shortcomings are my own responsibility.

Certain parts of this paper were previously published as part of Trask (1994-95).

References

Abbreviations:

ASJU	*: Anuario del Seminario de Filología Vasca "Julio de Urquijo"*
BRSVAP	*: Boletín de la Real Sociedad Vascongada de los Amigos del País*
FLV	*: Fontes Linguae Vasconum*
WIEV	*: Revista Internacional de Estudios Vascos*

Agud, Manuel and Antonio Tovar (1988-). *Materiales para un Diccionario Etimológico de la Lengua Vasca.* Published in fascicles in *ASJU*, XXII-l-present. Larger fascicles also available separately as *Diccionario etimológico vasco*, Donostia/San Sebastián: Seminario de Filología Vasca "Julio de Urquijo".

Aulestia, Gorka (1989). *Basque-English Dictionary.* Reno, NV: University of Nevada Press.

Azkue, Resurrección Maria de (1905). *Diccionario Vasco-Español-Francés* (2 vols.). Bilbao: privately published. Reprinted (1969) with a supplement by Lino de Aquesolo, Bilbao: La Gran Enciclopedia Vasca. Reprinted (1984) incorporating Michelena (1970a) as an introduction, Bilbao.

Bengtson, John D. (1990). 'An end to splendid isolation: the Macro-Caucasian phylum'. *Mother Tongue* 10.

——— (199la). 'Notes on Sino-Caucasian'. In Shevoroshkin (1991), pp. 67-129.

——— (199lb). 'Some Macro-Caucasian etymologies' . In Shevoroshkin (1991), pp. 130-141.

——— (199lc). 'Macro-Caucasian phonology (part I). In Shevoroshkin (1991), pp. 142-161.

——— (199ld). 'Macro-Caucasian: a historical linguistic hypothesis'. In Shevoroshkin (1991), pp. 162-170.

——— (199le). 'Some Sino-Caucasian etymologies'. In Shevoroshkin (1991), p. 172.

——— (199lf). 'Macro-Caucasian again'. *Mother Tongue* 13:19-26.

——— (1993). 'The Macro-Caucasic hypothesis'. *Dhumbadji!* 1(2): 3-6.

——— (1994a). 'Comment on Colarusso 1994'. *Mother Tongue* 22:13-16.

——— (1994b). 'On the genetic classification of Basque'. *Mother Tongue* 22:31-36.

——— (forthcoming). 'Is Basque isolated?'

Bengtson, John D. and Merritt Ruhlen (1994). 'Global etymologies'. In Ruhlen (1994a), pp. 277-336.

Blažek, Václav (1991). 'Basque and North Caucasian or Afroasiatic?'. *Mother Tongue* 14.

——— (1992). 'Materials for global etymologies'. In V. Shevoroshkin (ed.) (1992), *Reconstructing Languages and Cultures*, Bochum: Brockmeyer, pp. 37-40.

Blažek, Václav and John D. Bengtson (1995). 'Lexica Dene-Caucasica'. *Central Asiatic Journal* (forthcoming).

Bouda, Karl (1948). 'Baskisch und Kaukasisch'. *Zeitschrift für Phonetik* 2:182-202.

——— (1949). *Baskisch-Kaukasische Etymologien.* Heidelberg: Carl Winter.

——— (1951). 'L'euskaro-caucasique', in *Homenaje a D. Julio de Urquijo*, vol. III, pp. 207-232, San Sebastián.

——— (1952). 'Neue baskisch-kaukasische Etymologien', *Acta Salmanticensia Filosofía y Letras* 5(4):3-16.

Catford, J. C. (1991). 'The classification of Caucasian languages'. In Lamb and Mitchell (1991), pp. 232-268.

Čirikba, Vjačeslav A. (1985). 'Baskskij i severokavkazskie jazyki' ['Basque and the North Caucasian languages'], in B. B. Pietrovskij (ed.), *Drevnjaja Anatolija* [*Ancient Anatolia*], pp. 95-105. Moscow: Nauka.

Corominas, Joan (1969). 'Hurgando en los nombres de parentesco'. *FLV* 2(5):169-182.

Corominas, Joan and José A. Pascual (1980). *Diccionario Crítico Etimológico Castellano e Hispánico*, 6 vols. Madrid: Gredos.

Dumézil, Georges (1933). *Introduction à la grammaire comparée des langues caucasiennes du nord*, Paris: Bibliothèque de l'Institut Français de Leningrad, vol. XIV. Ch. 5: Caucasien du nord et basque, pp. 123-149.

Etxezarreta, Ramón (1983). *Hiztegi erotikoa.* Donostia: Hordago.

Gorrochategui, Joaquín (1987). 'Vasco-Celtica'. *ASJU* XXI(3):951-959.

Greenberg, Joseph H. (1987). *Language in the Americas.* Stanford: Stanford University Press.

Hualde, J. I., J. A. Lakarra, and R. L. Trask (eds), (forthcoming). *On the History of the Basque Language: Studies in Basque Historical Linguistics* [provisional title]. Amsterdam: John Benjamins. To appear 1995.

Hymes, Dell (1987). Review of Greenberg (1987). *Current Anthropology* 28:659-662.

Jacobson, William H., Jr. (1975). 'Historical implications of the western Basque tonal accent'. Ms., University of Nevada at Reno.

Kaiser, Mark and Vitaly Shevoroshkin (1988). 'Nostratic'. *Annual Review of Anthropology* 17:309-329.

Lafon, René (1933). 'Basque et langues kartvèles: À propos des postpositions basques formées au moyen de *-gan'*. *RIEV* 24:150-172.

——— (1944). *Le système du verbe basque au XVIe siècle*, 2 vols. Bordeaux. Reprinted (1980) in one volume, San Sebastián: ELKAR. Vol. 1: Appendice: 'Quelques correspondances morphologiques entre le basque et les langues caucasiques', pp. 527-535.

——— (1948). 'Correspondances basques-caucasiques'. *Eusko-Jakintza* 2:359-370.

——— (1951). 'Concordances morphologiques entre le basque et les langues caucasiques', part 1. *Word* 7:227-244.

——— (1952a). 'Le basque et les langues caucasiques (suite et fin)'. *Word* 8:80-94.

——— (1952b). 'Études basques et caucasiques'. *Acta Salmanticensia Filosofia y Letras* 5(2):5-91.

——— (1957). 'Le géorgien et le basque sont-ils des langues parentés?' *Bedi Kartlisa* 28:26-27.

——— (1967). 'La linguistique basque et caucasique'. *Revue de l'enseignement supérieur* (Paris) 3-4:56-66.

——— (1968). 'Pour la comparaison du basque et des langues caucasiques. I. Les indices de personne dans la conjugaison basque'. *Bedi Kartlisa* 28:13-26.

——— (1973). 'La langue basque', *Bulletin du Musée Basque* 38 [no. 60]:57- 120.

——— (1976). *Baskskij jazyk i basksko-kavkazskaja gipoteza* [*The Basque Language and the Basque-Caucasian Hypothesis*]. Tbilisi.

Lamb, Sydney M. and E. Douglas Mitchell (eds.) (1991). *Sprung from Some Common Source: Investigations into the Prehistory of Languages*. Stanford: Stanford University Press.

Lhande, Pierre (1926). *Dictionnaire Basque-Français*. Paris: Gabriel Beauchesne.

Meyer-Lübke, W. (1935). *Romanisches Etymologisches Wörterbuch*, 3rd ed. Heidelberg: Carl Winters Universitätsbuchhandlung.

Michelena, Luis (1949a). 'Notas de gramática histórica vasca'. In *Homenaje a D. Julio de Urquijo,* vol. 11, pp. 483-487. Reprinted in Michelena (1988), vol. I, pp. 445-447.

——— (1949b). 'Voces vascas'. *Emerita* 17:195-211. Reprinted in Michelena (1988), vol. I, pp. 448-457.

——— (1950a). 'Notas etimológicas vascas'. *Emerita* 18:467-481. Reprinted in Michelena (1988), vol. I, pp. 458-466.

——— (1950b). 'De etimología vasca'. *Emerita* 18: 193-203. Reprinted in Michelena (1988), vol. I, pp. 439-4

——— (1950c) . 'De fonética vasca. La aspiración intervocálica'. *BRSVAP* 6:443-459. Reprinted in Michelena (1988), vol. I, pp. 190-202.

——— (1953). Review of Bouda (1952). *BRSVAP* 9:141-144.

——— (1954). 'De onomástica aquitana'. *Pirineos* 10:409-458. Reprinted in Michelena (1985), pp. 409-445.

———(1955). 'Cuestiones relacionadas con la escritura ibérica'. *Emerita* 23:265-284. Reprinted in Michelena (1985), pp. 357-370.

——— (1956). 'Algunes nombres vascos de animales'. In *Homenaje a D. Joaquín Mendizabal, conde de Peñaflorida*, San Sebastián, pp. 313-320. Reprinted in Michelena (1988), vol. II, pp. 509-515.

——— (1957a). 'Las antiguas consonantes vascas'. *Miscelánea homenaje a A. Martinet,* vol. I, pp. 113-157, Universidad de La Laguna. Reprinted in Michelena (1988), vol. I, pp. 166-189. English translation by R. L. Trask, 'The ancient Basque consonants', in Hualde *et al* (forthcoming) .

——— (1957b). 'Basque et roman'. *Via Domitia* 4:12-25. Reprinted in Michelena (1988), vol. I, pp. 107-115.

——— (1961). 'Los nombres indígenas de la inscripción hispano-romano de Lerga (Navarra)'. *Príncipe de Viana* 22:65-74. Reprinted in Michelena (1985), pp. 446-457.

——— (1964a). *Textos arcaicos vascos.* Madrid: Minotauro. Second printing (1990).

——— (1964b). *Sobre el pasado de la lengua vasca.* San Sebastián: Auñamendi. Reprinted in Michelena (1988), vol. I, pp. 1-73.

——— (1967). Review of J. Hubschmid (1967), *Thesaurus Praeromanicus* II. *Zeitschrift für Romanische Philologie* 83:602-609. Reprinted in Michelena (1985), pp. 321-328.

Michelena, Luis (1968). 'L'euskaro-caucasien'. In A. Martinet (ed.), *Le langage*, Paris, pp. 1414-1437. Reprinted in Michelena (1985), pp. 458-475.

——— (1970a). *Estudio sobre las fuentes del diccionario de Azkue.* Bilbao: Publicaciones del Centro de Estudios Históricos de Vizcaya. Reprinted in the (1984) printing of Azkue (1905), pp. 17-104.

——— (1970b). 'Nombre y verbo en la etimología vasca'. *FLV* 2: 67-93. Reprinted in Michelena (1987), pp. 283-309.

——— (1971a). 'Descubrimiento y redescubrimiento en textos vascos'. *FLV* 3:148-169. Reprinted in Michelena (1988), vol. II, pp. 713-728.

——— (1971b). 'Toponimia, léxico y gramática'. *FLV* 3:241-267. Reprinted in Michelena (1987), pp. 141-167.

——— (1971c) 'Egunak eta egun-izenak'. *Munibe* 23:583-591. Reprinted in Michelena (1987), pp. 269-282. Also reprinted in K. Mitxelena (1988), *Euskal Idazlan Hautatuak*, P. Altuna (ed.), Donostia, pp. 93-110.

——— (1972). 'Léxico vasco y etimología'. In *Beiträge zur Romanistik und allgemeinen Sprachwissenschaft: Festschrift Wilhelm Giese*, Hamburg, pp. 79-95. Reprinted in Michelena (1987), pp. 337-348.

——— (1973). *Apellidos vascos,* 3rd ed. [lst ed. 1953] San Sebastián: Txertoa.

——— (1974). 'El elemento latino-románico en la lengua vasca'. *FLV* 6:183-209. Reprinted in Michelena (1987), pp. 195-219.

——— (1977). *Fonética histórica vasca,* 2nd (corrected and expanded) edition. [lst edition 1961; 3rd edition 1985] San Sebastián: Publicaciones del Seminario Julio de Urquijo de la Excma. Diputación Provincial de Guipúzocoa.

——— (1985). *Lengua e historia.* Madrid: Paraninfo.

———(1987). *Palabras y textos.* Vitoria.

——— (1988). *Sobre historia de la lengua vasca* (2 vols), ed. by J. A. Lakarra. Donostia-San Sebastián: Anejos del *Anuario del Seminario de Filología Vasca "Julio de Urquijo"* 10.

Michelena, Luis, J. Caro and Antonio Tovar (1966). 'La obra lexicográfica de don R. Mª. de Azkue'. In *Don R. Mª. de Azkue lexicógrafo, folklorista y gramático.*

Ruhlen, Merritt (1994a). *On the Origin of Languages: Studies in Linguistic Taxonomy.* Stanford: Stanford University Press.

——— (1994b). 'An overview of genetic classification'. In Ruhlen (199la), pp. 9-38.

Salmons, Joe (1992). 'A look at the data for a global etymology: **tik* "finger"'. In G. W. Davis and G. K. Iverson (eds) (1992), *Explanation in Historical Linguistics,* Amsterdam: John Benjamins, pp. 207-228.

Schuchardt, Hugo (1913). 'Baskische-hamitische Wortvergleichungen'. *RIEV* VII:309ff.

——— (1923). *Primitiae Linguae Vasconum.* Halle. Reprinted (1968) with an introduction and a bibliography by Antonio Tovar, Tübingen: Max Niemeyer Verlag.

Shevoroshkin, Vitaly (ed.) (1991). *Dene-Sino-Caucasian Languages.* Bochum: Brockmeyer.

Shevoroshkin, Vitaly and Alexis Manaster Ramer (1991). 'Some recent work on the remote relations of languages'. In Lamb and Mitchell (1991), pp. 178-203.

Starostin, Sergei A. and Merritt Ruhlen (1994). 'Proto-Yeniseian reconstructions, with extra-Yeniseian comparisons'. In Ruhlen (1994a), pp. 70-92.

Tovar, Antonio (1950). *La lengua vasca.* San Sebastián : Biblioteca Vascongada de los Amigos del País.

——— (1958). Article in *Festschrift W. von Wartburg,* 831ff., cited in Agud and Tovar (1988-) under *amar.*

——— (1959). *El euskera y sus parientes.* Madrid.

Tovar, Antonio, K. Bouda, R. Lafon, L. Michelena, W. Vycichl, and M. Swadesh (1961). 'El método lexico-estadístico y su aplicación a las relaciones de vascuence'. *BRSVAP* 17:249-281.

Trask, R. L. (1977). 'Historical syntax and Basque verbal morphology: two hypotheses'. In W. A. Douglass, R. W. Etulain, and W. H. Jacobsen, Jr. (eds), *Anglo-American Contributions to Basque Studies: Essays in Honor of Jon Bilbao*, Reno, NV: Desert Research Institute Publications on the Social Sciences No. 13, pp. 203-217.

——— (1981). 'Basque verbal morphology'. *Iker-1: Euskalarien Nazioarteko Jardunaldiak*, Bilbao: Euskaltzaindia, pp. 285-304.

——— (1985). 'On the reconstruction of pre-Basque phonology'. In J. L. Melena (ed.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae*, Vitoria: Instituto de Ciencias de la Antigüedad, vol. 2, pp. 885-891.

——— (1990). 'The *-n* class of verbs in Basque'. *Transactions of the Philological Society* 88:111-128.

Trask, R. L. (1994-95). 'Basque: the search for relatives (part 1)'. *Dhumbadji!* 2(1):3-54.

——— (forthcoming a). 'On the history of the non-finite verb forms in Basque'. In Hualde *et al* (forthcoming).

——— (forthcoming b). 'Origin and relatives of the Basque language: review of the evidence'. In Hualde et al (forthcoming).

——— (forthcoming c). *The History of Basque.* London: Routledge. To appear 1996.

Trombetti, Alfredo (1925). *Origini della lingua basca.* Bologna: Memorie della R. Accademia delle Scienze dell' Istituto di Bologna.

Uhlenbeck, C. C. (1923). *Over een mogelijke verwantschap van het Baskisch met de palaeo-kaukasische talen.* Amsterdam.

——— (1924). 'De la possibilité d'une parenté entre le basque et les langues caucasiques'. *RIEV* XV:565-588.

——— (1927). 'Die mit *b-* anlautenden Korperteilnamen des Baskischen'. In *Festschrift Meinhof.* Hamburg.

——— (1940-41). 'Vorlateinische indogermanische Anklänge im Baskischen'. *Anthropos* 35-36:202ff.

——— (1946). *Gestaafde en vermeende affiniteiten van het Baskisch.* Amsterdam. French translation (1947), 'Affinités prouvées et presumées de la lange basque'. *RIEV* I:171ff.

——— (1947). 'La langue basque et la linguistique générale'. *Lingua* 1:59-76.

Vogt, Hans (1942). 'La parenté des langues caucasiques'. *Norsk Tidskrift for Sprogvidenskap* 9:321-338.

——— (1955). 'Le basque et les langues caucasiques'. *Bulletin de la Société Linguistique* 51:121-147.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

R . L. TRASK, BASQUE AND DENE-CAUCASIAN: A CRITIQUE FROM THE BASQUE SIDE

Paul K. Benedict

[Editor's Note: Paul Benedict is a founding father of ASLIP and a long ranger for many decades. His expertise on Southeast Asian languages, especially Sino-Tibetan and Daic, is well-known. Since he construed his task as limited to Sino-Tibetan, his comment is very short.]

1. Sino-Tibetan roots not cited. In view of the fact that Sino-Tibetan has been viewed as a member of Dene-Caucasic, however, it appears to be noteworthy that the lack of comparable forms in Sino-Tibetan is even greater than might have been anticipated by chance![1] Actually, only a pair of comparable forms has been noted, one (213) with "nursery" ("universal") vibes, the other (269) with "imitative". Shouldn't one have done better than that by chance? I did find one nice look-alike, this for Basque *tipi* 'small' cited under (272), but it is from Austro-Thai, viz., PAT **tipits* 'small' (see Glossary in my *Japanese/Austro-Tai* book). Want to start an Austro-Tai/Dene-Caucasic hypothesis, anyone?

2. Re *koin* 'dog'. All good Dene-Caucasianists should realize that the Proto-Sino-Tibetan (PST) root here is **kwəy* (tone *B), widespread in Tibeto-Burman, represented in Chinese by a form with final *-n* standing for the PST "collective plural" suffixed **-n* (see my *Sino-Tibetan* book). The citation of the above as an **-n* final root is hereafter *verboten.*

3. One over-all criticism: the failure to stick with Earliest Possible Forms, i.e., to cite only roots, modifying them as might be required (but never without very explicit basis for so doing), or, in the absence of such, citing along with the present form what the ancestral form (root) would be, this often requiring complex structures, e.g., Written Tibetan *khyi* 'dog' < WTB- level (or simply WTB-) **k(w)[i/ə]y*.

Basque has sufficient dialectical variation to require the setting up of a Proto-Basque, with the help of the available historical information. I — and I am sure others — would like to see this done at this stage before going further. Can do? Anyone?

[1] [Editor's Note: The reader is reminded that the proposed etymologies linking Basque and Dene-Caucasic — which we are dealing with in Trask's paper — are those presented by Trask because they involve Basque. There is a larger set of proposed etymologies involving Sino-Tibetan in a major way with Na-Dene, Yeneseian, and Caucasic which were not included in the scope of Trask's paper. Those involving Sino-Tibetan and Caucasic have been the central core of Starostin's hypothesis — Sino-Caucasian.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

BASQUE: AN ORPHAN FOREVER? A RESPONSE TO TRASK

John D. Bengtson
Minneapolis, MN

1. Of Errors and "Errors"

Trask is a specialist in Basque, and, as such, I have respect for his knowledge of the facts of the modern language and attested forms. At the outset, let me state that I admit that Trask is correct about the following facts:

1. Basque *atal* does not mean 'limb', but 'segment, fragment, portion'.
2. Basque *bargo* 'young pig' is not directly related to Caucasic **wārtʼwə* 'pig'. (The phonologically regular Basque cognate is probably *urde* 'pig, swine": see "Vasco-Caucasic Phonology," below.)
3. Basque *bekoki* 'forehead' is better analyzed as *beko-ki* rather than my earlier **be-koki*.
4. Basque *gela* 'room, chamber' is a loan from Latin *cella*, thus not directly related to Caucasic **qəlV* 'house'.
5. Basque *tiña* 'ringworm' is a loan from Latin/Romance *tinea* and not directly connected with Caucasic **tʼānhV* 'nit'.
6. I now agree with Trask's analysis of Basque *makutsik* 'in one's shirtsleeves'.
7. Trask seems to be correct that Basque **maño* 'masculine' does not exist. It is probably a variant of *mando* 'mule' (Spanish *macho* 'masculine, mule').
8. Basque *paru* ~ *pau* 'pole' is a loanword from Latin/Romance *palu* and thus not connected with Burushaski *-pawo* 'stick, staff'.
9. Basque *santan* 'ever', *senton* 'old man', etc., are clearly loans from Romance and thus not connected with Caucasic **śüno* 'year', etc.

There may be a handful of other cases where Trask's extensive knowledge of Basque succeeds in "destroying" (as he is so fond of saying) Vasco-Caucasic comparisons, and we are grateful to the extent that this information genuinely corrects errors due to ignorance.

But corrections of this type make up a small fraction of Trask critique. The great majority of "errors" attributed to me and other vasco-caucasologists are simply cases where we disagree with (a) Trask's "Pre-Basque Phonology," and (b) Trask's (and/or Michelena's) pet etymological solutions (usually supposing Basque borrowings from a vague "Romance").

So, our "errors," for the most part, consist of our reluctance to accept, on faith, it seems, Trask's and Michelena's *hypotheses* about prehistoric Basque phonology and specific word origins. Hypotheses is the operative word here, since that is exactly what "Pre-Basque Phonology" (see "Pre-Basque Phonology: A Hypothesis Built on Sand," below), and what we consider fantastical Latinate etymologies (see "Is It Romance?," below) are, though Trask almost always presents them *as if they were established facts.*

But even if we accept all of Trask's alleged "errors" as actual errors (which I emphatically do not), would that invalidate the Vasco-Caucasic and Dene-Caucasic hypotheses and prove that our evidence "amounts to precisely zero," as Trask contends? Here, I shall remind the reader of the early history of Indo-European linguistics (cf. Pedersen 1962:242, 254, 255):

> The insufficient basis in method explains many of the mistakes made by the first students of comparative linguistics [e.g., Rask, Bopp, Grimm] ... Rask's *Undersøgelse* has its obvious faults ... mistakes he made both in the phonology and in the comparison of inflectional endings ... Bopp treats Persian and Germanic under one heading. Certainly his eye for relationships was none too keen here. He allowed himself to be deluded by certain superficial resemblances ... [etc.]

Yet no one today claims that Rask's, Bopp's, and Grimm's work was "worthless" (Trask 1994-1995:43, referring to Čirikba), that they were "total failures" (Trask 1994-1995:48), or that their evidence for Indo-European "amounts to precisely zero" (Trask 1995)! To the contrary, they are rightly and universally recognized as the founders of Indo-European comparative linguistics.

The French vasconist Michel Morvan, while agreeing in some details with Trask, nevertheless concludes that (Morvan 1992:365):

> ... the work of the omni-comparativists is necessary, and those who will not hear of it are wrong. But the errors are numerous and inevitable. One does not progress without them ...

Earlier in the same article, Morvan (1992:357) warns of

> ... the widespread trap to which many linguists succumb, which, from fear of making an error, and perhaps from too much training, is translated into an excess of caution, which is itself as disastrous as the lack of caution. There can no longer be any doubt that there is some truth in linguistic superfamilies such as Nostratic or Dene-Caucasian.

The present state of Dene-Caucasic studies is comparable to that of Indo-European studies at the time of Rask, Bopp, and Grimm. Almost everything remains to be done, and so few workers!

Trask's Errors: Since Trask has devoted a great deal of space, time, and energy to listing the errors (and "errors") committed by me and other vasco-caucasologists, it seems only right that I return the favor and take a much smaller space here to list his own errors, as I see them. These errors fall under two headings: (1) errors of fact, and (2) errors of method.

Errors of Fact: (1) Trask claims that "roots in North Caucasian languages are typically very short, often no more than a single consonant," for which no reference is given. Trask must be referring here only to *West* (= Northwest) Caucasic, where it is true that most but not all roots are of the type CV (but not C!). The latter "vowelless hypothesis" refers only to Kabardian, but is thought to be untenable by Catford (1981:248). For Proto-East Caucasic and Proto-Caucasic itself, the root structure is very commonly more complex: CVCV, and other types such as CVRCV and CVCVCV (see Table 1 for examples). Trask may not like it, but clearly these root structures are highly compatible with those of Basque.

Related to this claim is the statement that "the first attempt at reconstructing Proto-North-Caucasian has provisionally set up the astounding total of 180 consonants." While this was true at one time, "recent developments in the field of North Caucasian comparative phonology ... have now made it possible to modify and simplify this system considerably" (Starostin 1989:47). In the following pages (48-49) of the same article, Starostin's revised table of Dene-Caucasic correspondences lists only 42 proto-segments (12 of which appear to be clusters), which in turn correspond (due to mergers) with *only 32* Caucasic phonemes. I find this inventory highly compatible with that of Basque.

(2) Trask further claims that "All of [the vasco-caucasologists] appear to be entirely unaware of the existence of Michelena's [*Fonética histórica vasca*]." As can be seen from one of my published papers (Bengtson 1992), I have been aware of Michelena's phonological hypotheses for several years. While finding the book very useful as a mine of the phonological facts of Basque, I found I could not go along with his speculations about "Pre-Basque Phonology" for reasons discussed in detail below.

(3) In his attempt to show that all instances of Basque *h*'s are not etymological, Trask claims that the Zuberoan (Souletin) dialect of Basque "regularly extends the aspiration to vowel-initial monosyllables," thus *hor* 'dog', *huñ* 'foot', etc., derive from *or*, *oin*, etc., according to Trask. I would counter that assertion with the fact that in some monosyllables the aspiration contrast is still maintained, e.g.:

(Zuberoan) *har* 'worm' vs. *ar* 'male'
hots 'come on' vs. *ots* 'male (animal)'
hütz 'fart' vs. *ütz* 'to leave'
(Larrasquet 1939).

(4) In several places, Trask makes the pronouncement that "Basque *h* is never etymological." (How Trask, with no external comparison to depend on, can extend this and other phonological "facts" indefinitely into the prehistory of Basque is never adequately explained.) In any event, let us see what light external comparison sheds on this question:

Basque *hor* 'dog' : cf. Caucasic: Budukh *Xor* 'dog', etc. (see Table 1);
Basque *hauts* 'dust, powder' : cf. Caucasic **XurtV* 'foam, scum'; Burushaski *xurc* 'dust';
Basque *hari, hal-* 'thread' : cf. Caucasic **XalV* (Chechen *Xal* 'a piece of thread'; Tsez *Xero* 'sinew'; Lezgi *ɣal* 'thread', etc.);
Basque *ahari* 'sheep' : cf. Caucasic **X[ə]rV* (Andi *Xor-ol* 'sheep, sheep-flock'; Hunzib *Xor* 'ram', etc.);
Basque *alha* (< **hala*) 'pasture, grassland, to graze' : cf. Caucasic **XalV* (Chechen *Xal* 'stalk'; Lak *χ̄ala* 'stalk, grass', *χ̄ulu* 'hay', etc.);
Basque *herri* 'country, town, inhabited place, people' : cf. Caucasic **XwōrV* 'village' (Avar *Xur*; Lezgi *Xür*; etc.).

Here in six cases, Basque *h* corresponds to Caucasic *X* (unvoiced uvular fricative), so it would seem that *h* did not spring up *ex nihilo* as Trask would have it. Note further:

Basque *har* 'worm' : cf. Caucasic: Avar *ħapára* 'worm', etc. (see Table 2);
Basque *hur* 'hazelnut' : cf. Caucasic **ħwort'V* (Hunzib *het'e* 'walnut'; Chechen *bʕara* 'hazelnut'; Abkhaz *a-ra* 'nut', etc.);
Basque *habe* 'pillar, beam' : cf. Caucasic **ħwVbV* (Avar *ħubi* 'pillar, pole, trunk, stem'; Hunzib *hebo* 'stick', etc.);
Basque *hur* 'water' : cf. Caucasic: Avar *ħor* 'lake', etc. (see Table 1).

Here we have four cases where Basque *h* corresponds to Caucasic *ħ* (unvoiced pharyngal fricative), which at once refutes Trask's claim that *all h*'s are unetymological, and also solves his problem with the disparity between the Basque and Caucasic consonant systems. The solution is that Basque *h* represents a *phonemic merger* of several phonemes: a uvular *X*, pharyngal *ħ*, and probably also the usual glottal *h*, and the aspirated velar k^h of Proto-Vasco-Caucasic.

(5) In several cases, Trask attempts to "destroy" etymologies by postulating an internal *-n-*, as in the cases of *azeri* 'fox' and *ziho* 'tallow'. Trask tries to derive *izotz* from *ihintz-hotz* 'dew-cold'.

But this kind of change leaves a trace in the Zuberoa dialect: the vowels of the word are nasalized, thus, *uhure* [ũhũẽ] from Latin *honore-*; *ihi* [ĩhĩ] 'reed, rush' also bears witness to an earlier nasal in the word. But *axeri* [ašéy] 'fox', *ziho* [sĩho] 'tallow', and *izotz* [ĩsõc] 'frost' have no nasal vowels (Larrasquet 1939), so Trask's claims fall apart.

Errors of Method: (1) Arbitrary segmentation: Trask accuses me and the other vasco-caucasologists of "arbitrary segmentation." But taking a closer look, we see that Vasco-Caucasic segmentations are almost always simply the segregation of fossilized class prefixes, e.g., *i-zar* 'star', *bi-hotz* 'heart', *u-kondo* 'elbow'. On the other hand, some of the segmentations Trask accepts and advocates seem very odd, for example, Blažek and I propose:

Basque *astigar* 'maple' : cf. Caucasic: Batsbi *stagar* (not **stager*), Chechen *stajr* 'Acer platanoides'; cf. Hurrian **taskar-* 'boxwood'. (The only segmentation we make is to separate the prefix *a-*.)

But Trask prefers the segmentation *gazt-igar*, based on a localized form, thus a combination of Romance *cast-* (an element in three names, according to Trask) and Basque *ihar* 'maple'. But, in order to get *cast-*, he has to further segment its source, Latin *castanea* 'chestnut tree', which became Basque *gaztaiña* 'chestnut'. My interpretation of dialect geography indicates that *gaztigar* (LH, L, R) is a central innovation arising from contamination with *gaztaiña*, because the older *astigar*, *aztigar* persist at the extremes of Basque country (B, Z). The Basque-Nakh comparison is straightforward, indeed almost exact phonetically and semantically. But Trask will go to absurd lengths to avoid such facts.

Trask has also criticized my etymology of Basque *ukondo* 'elbow' as "arbitrary segmentation":

> Basque *ukondo* 'elbow': cf. Caucasic: Lezgi *q'ünt* 'elbow'; Tsez *q'o̧ntu* 'knee'; Khwarshi *q'ontu* 'knee', etc. (Proto-Caucasic **q'HwəntV*). (Again, the only segmentation is that of the fossilized class prefix *u-*.)

But rather than accept this phonetically and semantically perfect comparison, Trask would rather cut the word up into **uko* (a hypothetical Basque word for 'hand, arm') plus *ondo* 'bottom', which he claims is a loan word from Romance. Thus again, we have a Basque-Latin hybrid word, though this would be unexpected in words for 'elbow' or 'knee', which are typically highly stable. No other European language, as far as I know, has a word for 'elbow' that is a half-native / half-foreign compound of this type. As so often with Trask, Basque is exempt from the usual rules of comparative linguistics, because for him there is nothing to compare Basque with!

Another example is the Basque word for 'gall, bile', where Trask again disputes the segmentation proposed by Bouda and Chirikba, and to which I assent:

> Basque *beazun* (G) = **be-a-sūn* : cf. Caucasic: Avar *c̄in*
> *beaztun* (B) = **be-a-stūn* : cf. Caucasic: Chechen *stim*
> *beazuma* (L, Z) = **be-a-sūma* : cf. Caucasic: Dargi *sume ~ time ~ himi*;
> (Proto-Caucasic **cwă̈jme* 'gall, anger').

The element *be-* is the remnant of a fossilized class marker (stage III article) that is found in a number of Basque words, and constitutes some of the strongest morphological evidence for Vasco-Caucasic. The element *-a-* is still mysterious to me, but the third element, variously *-ztun*, *-zun*, *-zuma* we believe to be cognate with the Caucasic words listed above. Note that the Basque dialectal variants recapitulate some of the Caucasic local variants.

But Trask does not agree: he insists on something like **be(h)a(t)z-dun* or **beCa(t)z-dun*, where the first element is the root and the second is the "derivational suffix *-dun* 'having'." But this does not explain the northern *-zuma* or even the more general *-zun*, though both have clear Caucasic parallels. Trask can give us neither origin nor external parallels for his supposed **baCa(t)z-* — it is simply a lexical ghost, spun out of thin air, like so many of the traditional vasconist etymologies.

We see again and again that Trask will go to almost any lengths to explain away or "destroy" Vasco-Caucasic etymologies that otherwise look semantically and phonetically plausible to most linguists. In doing so he makes his own bizarre segmentations, making Basque a patchwork of "Romance" elements combined willy-nilly with "Basque" ghost words for which Trask can provide no origin or external confirmation.

All etymologists must cut words into segments sometimes. As I have just shown, the way the cuts are made depend on the author's hypothesis.

(2) Phonological errors: (a) "Insertion" of *h* and *g*: Trask, in his discussion of Basque words for 'tomorrow', contends that "*h* is not etymological, and western *bigar*" shows "the common western insertion of *g* to separate vowels in hiatus." So for Trask, apparently, the hiatus form *biar* comes first, and the forms *bihar* and *bigar* [biɣaɾ̄] are secondary. But this is totally the reverse of the usual comparative method, which would assume that the forms with an intervocalic obstruent represent the earlier stage. Here is my solution: early Basque **bixar* had [x] (a voiceless velar fricative), which was voiced to *bigar* (G), phonetically [biɣaɾ̄], in parts of Gipuzkoa, but elsewhere

[x] merged with *h* (see above), and the *h* in turn became mute in some dialects. So the actual development was the *reverse* of Trask's scenario. The external parallels are Caucasic **bəgə* 'evening, morning, tomorrow' (see Table 1), and more distantly Yeniseian **pVk-* 'morning' (Ket *hɨgem*, Pumpokol *-paga*) (Starostin and Ruhlen 1994: 82), one of the clearest Dene-Caucasic etymologies.

Similarly, Trask attacks the Dene-Caucasic etymology of Basque *aho* 'mouth', contending that "Bengtson's **axʷo* is no more than a flight of fancy." In point of fact, I was not the author of this reconstruction: it is the work of Václav Blažek (my co-author of "Lexica Dene-Caucasica"), and it is a brilliant stroke, accounting for all the local Basque variants:

Basque **axʷo* > **aɣʷo* > *abo* [aβo] (B Arratis, Orozko)

~ **axo* < *aho* (LN, L, Z) > *ao* (HN, G, B, R); **aɣo* = *ago* (HN, B, R, Z)

Blažek's reconstruction accounts for all the Basque forms, but Trask, with his position that *h* is never etymological, is forced to posit *ao* as the original form, which is again the reverse of the usual comparative method.

(b) Phonetic inconsistency: There is no phonological consistency to many of Trask's so-called "Romance" loanwords. For example, he derives *horma* 'ice' from Latin *forma* 'form' (see below for semantic latitude), disdaining my comparison with Burushaski *ɣamu* 'ice'. On the other hand, Trask claims that Basque *ondo* 'bottom', etc. is derived from Latin *fundu*. Should we not then have *horma* and **hondo* in the aspirating dialects? To the contrary, they have *horma* but *ondo/undo*.

The comparison Basque *horma* 'ice': Burushaski *ɣamu* 'ice' (where *ɣ* = (= a voiced *uvular* fricative; see Tiffou and Pesot 1989:8ff) is semantically exact, and phonetically the correspondence of Basque *h* to Burushaski *ɣ* is confirmed by other comparisons (a fact Trask ignores in his critique of my Vasco-Caucasic phonology):

Basque *huñ* (Z) 'foot' : Burushaski *-ɣán* 'heel'
Basque *hausin* (LN, L, Z) 'nettle': Burushaski *ɣašu* 'nettle'
Basque *ohoin* 'thief' : Burushaski *ɣin* 'thief'
Basque *harri* 'stone' : Burushaski *ɣoro* 'stones'
Basque *hümoi* (Z) 'womb' : Burushaski *-ɣúmur* 'bowels, insides'

These comparisons provide further refutation of Trask's claim that Basque *h* came out of nothing. As usual, Trask cannot accept such a clear and simple etymology, so instead he prefers to derive *horma* 'ice' from Latin *forma* 'form'!

The so-called Romance derivation of *ondo* can also be called into question. Besides the phonetic problem, the word appears to be old in Basque, with meanings as varied as 'side; stem, trunk', and it appears in many compounds, e.g., in Zuberoan: *zuhañ-ondo* 'foot of a tree', *bazkal-ondo* 'after breakfast', *ondo-amen* 'consequence'. Latin cannot explain all this, so I suggest instead a comparison with Caucasic **ʔəmdV* 'joint, bone', etc. (Nikolaev and Starostin 1992). The meaning 'joint' appears also in some Basque words, e.g., *gar-ondo* (HB, B, LN, G, L, R) 'neck'. I think the existence of a native Basque *ondo* is more plausible than its derivation from Latin, but of course in recent times the chance resemblance with Spanish *hondo* 'deep' may have influenced the Basque word semantically.

In short, Trask accepts inconsistent and improbable phonetic transformations because, with no external relationships, there is nothing to check them against. In contrast, the Vasco-Caucasic hypothesis offers an increasing number of phonetically and semantically clear comparisons which provide a solid basis for Basque etymology.

(3) Semantic latitude: In his zeal to keep Basque isolated, Trask accepts some highly improbable semantic equations. I have already mentioned 'ice: form' above. Others are 'dung : saffron-colored', 'marrow : rope', 'bad :

old', and 'side : trousers'! The respective Vasco-Caucasic parallels are all much better semantically, e.g., 'ice : ice, frost', 'dung : dung', 'marrow, brain, pith : blood, meat', 'bad : wild', 'side : side, cheek'. This semantic latitude is just another aspect of Trask's rule of Basque etymology, which seems to be "*Any* Basque-Latin or Basque-'Romance' comparison, no matter how semantically or phonetically improbable, is preferable to a clear and simple Vasco-Caucasic comparison."

(4) Disregard of dialect geography: Trask shows no comprehension of the principles of dialect geography, and seems to pick his 'original' Basque forms according to what fits his hypothesis of isolation. I have already discussed Basque *astigar* 'maple tree', where Trask picks the form *gaztigar* as original because then he can make it into a Latin-Basque compound and avoid accepting the clear comparison with Batsbi. Unfortunately for Trask, *gaztigar* is restricted to central Basque dialects, is clearly a recent contamination with *gaztaiña* 'chestnut tree', and the forms without *g-* are found in the outlying dialects, especially Bizkaia and Zuberoa, indicating their originality and the correctness of the Basque-Nakh comparison.

The importance of Bizkaia, in the far west, and Zuberoa, in the far northeast of Basque country, is shown again by the words for 'marrow, pith, brain': the *mun* forms are central (indicating contamination with a different root: *mami*), Bizkaian has *un* (from **hun*) 'médula, tuétano, seso' and Zuberoan has *hün* 'cerveau, moelle des os'. So these are the archaic forms, which I compare with Caucasic **hwĕʔnV* 'blood' ('meat' in Avar).

Another case is Basque *ukondo* 'elbow' (see above), which Trask seems to accept as a compound of *ukarai* 'wrist' and *ondo* (see above), ignoring the fact that *ukalondo* is another central Basque innovation, and Bizkaian *ukando*, *ukondo* is more archaic. So in virtually every case, a proper understanding of dialect geography confirms Vasco-Caucasic comparisons.

(5) Historically implausible etymologies and/or mode of transmission: In several cases Trask accepts derivations, usually from Latin, that make one wonder how the word got from here to there. Thus Trask dismisses Basque *kuma* 'mane, horsehair' as a "simple borrowing from Old Spanish *coma* 'horsehair, mane'. Yet Corominas and Pascual (1980), one of Trask's favorite dictionaries, points out that *coma* was rare, and "no se empleó más que como latinismo poético." The Latin *coma* was itself poetic and borrowed from Greek. So what was the mode of transmission from Latin to Basque? Are we to believe that the Latin and Spanish poets held seminars in Basque country, and that crusty Bizkaian shepherds and farmers suddenly gave up their native word for 'mane' and adopted the word used by the foreign poets? The variant *kima* (though only Gipuzkoan) causes further difficulty for the Spanish derivation. I still prefer the Vasco-Caucasic etymology, and the Latin parallel seems to be a chance resemblance.

Similarly Trask derives Basque *inguru* 'vicinity; around, near' from Latin *in gyrum* or *in gyru* 'in a circle', again ultimately of Greek origin. But how did it get from here to there? Likewise Basque *agure* 'old man', derived by Trask from the Latin vocative *avule* 'grandpa!' Apart from the phonetic problem, is there any evidence that Basque has borrowed Latin words *complete with the correct vocative ending*? (It reminds me of the purported derivation of Basque *abere*, *abel-* 'cattle' from the Latin *infinitive habere* 'to have', discussed below.) Again, there is no plausible mode of transmission. Trask forgets that linguistic borrowing requires a historically plausible bilingual community. Words do not just float from place to place.

(6) Errors of interpretation of evidence: For example, Trask, in attempting to dismiss the morphological evidence for Vasco-Caucasic, points out that many grammatical endings contain only one consonant. Well, of course; the same argument could be applied to Indo-European, but this is irrelevant. Any single element in isolation is not convincing evidence, but Trask forgets that it is the *cumulative* evidence that matters. It is morphological *patterns* (see Table 3) that prove Vasco-Caucasic and Dene-Caucasic, as shown above in several cases. There are lexical patterns as well: note that the Basque words for 'dog', 'wolf', and 'fox' *all* have Caucasic parallels:

Basque *hor* 'dog' : cf. Budukh *Xor* 'dog', etc. (Table 1)
Basque *otso* 'wolf' : cf. Andi *boc'o* 'wolf', etc.
Basque *azeri* 'fox' : cf. Tindi, Botlikh *sari* 'fox', etc.

Of course Trask cannot accept the third set here. Instead he concurs in Michelena's fantastic derivation of 'fox' from a Latin name *Asinarius*! All that is needed to "destroy" the latter derivation is to point out that there is no nasal

vowel in the Zuberoan form *axeri* [ašéy], as already noted above. Furthermore, the Bizkaian forms *azegari, azagari* are shown by dialect geography to be the most archaic. The fricative *g*, far from being "inserted" as Trask alleges, has a direct correspondence in the earliest Caucasic form proposed by Nikolaev and Starostin:

Basque (B) **a z e g a r i** 'fox'
Caucasic *** c E Hw ō l ē** 'fox'

I do not believe that Trask actually believes in these poor methods, nor would he teach them to his students. But it should be clear by now that his overriding zeal to keep Basque isolated causes him to propose things that his good sense would veto in other situations. Where Basque is concerned, virtually any explanation is acceptable to Trask, as long as it does not involve accepting a Caucasic or Burushaski parallel.

"Pre-Basque Phonology": A Hypothesis Built on Sand

As stated above, I have known about the Michelena-Trask "Pre-Basque Phonology" for several years, but do not accept it for the following reasons: (1) It is based on faulty logic: one possible interpretation of the evidence is adopted, ignoring other possible interpretations; (2) It is based in part on ancient and scanty inscriptions, of which there could again be multiple interpretations; (3) It ignores the testimony of modern dialects; (4) It is not supported by any external comparison, and in fact external comparison refutes most of it.

In more detail: (1) Michelena and Trask claim that because some old Basque loans show Basque *b- d- g-* from Latin *p- t- k-*, ancient Basque therefore had no voicing contrast in initial position. This is one possible interpretation, but not *the only* possible interpretation. Suppose for a moment that ancient Basque had the contrast of voiced stops (*b d g*) with voiceless *aspirated* stops (*ph th kh*) in initial position. The language of the Latin-speaking soldiers and merchants, on the other hand, had the contrast of voiced stops with voiceless *unaspirated* stops (*p t k*). Thus in words such as Latin *pice-* 'pitch', *turre* 'tower', *cella* 'room', the Basques would have no exactly corresponding initial consonant; Latin *p- t- c-* (*k-*) fell somewhere between the Basque *b- d- g-* and *ph- th- kh-*. While Michelena and Trask emphasize the *b d g* cases, there is in fact much vacillation between *b* and *p*, *d* and *t*, *g* and *k*:

Basque *bikhe* (LN, L), *bike* (L, HN), versus *pika* (B, G), *phike* (Z)
'pitch' (Latin *pice-*).
Basque *dorre* (HN, LN) versus *torre* (B, G, LN, R), *thorre* Z 'tower' (Latin *turre-*)
Basque *gerezi* (HN, LN, G, R, Z) versus *keriza* (B, G) 'cherry'
(Latin *ceresea*)

In spite of this vacillation, Michelena and Trask arbitrarily selected the *b d g* forms as original, and they have further generalized from this that "no native Basque word can begin with any *p, t, k*", etc. Now that this has become the rule, all such words are consigned to "non-native" status, regardless of any evidence to the contrary. I cannot think of a better example of illogical circularity!

(2) Much of "Pre-Basque Phonology" is based on forms from ancient Aquitanian inscriptions. It should hardly be necessary to enumerate the problems with this: the inscriptions are scanty, interpretation is sometimes doubtful, and the phonetic values are uncertain. For example, the Aquitanian name *Sembe* is invoked as evidence that Basque *seme* 'son' came from older **senbe*, and thus that there was no *m* in Pre-Basque! This seems to me to be mighty slim evidence for such sweeping pronouncements!

Some of the inscriptional evidence actually contradicts Trask, as when Aquitanian *h* shows the aspirate is ancient.

(3) It is contradicted by the evidence of modern Basque dialects: According to Hualde's (1991) excellent Basque phonology, which is based on facts, not speculation, a "common consonant inventory in Basque" consists of 23 segments (p. 10); other dialects discussed in the book have minimum inventories of 19-21 consonants. Trask's

claim that "Pre-Basque" had no more than 16n consonants, and possibly "as few as eight" (!) is not supported by extant forms of Basque, nor by any evidence at all! It is simply the result of the faulty logic discussed under (1).

(4) Finally, Trask can cite no external parallels to confirm his phonology, since for him there are no external parallels. The etymologies adduced in support of Casco-Caucasic in fact refute Trask's "phonology", showing, for example, that ancient Basque had a voiced/aspirated stop contrast (*b* : *ph*, *d* : *th*, *g* : *kh*), and that Basque *h* did not arise out of nothing, but represents a merger of several earlier phonemes (see above).

I therefore contend that there is *zero evidence* (as Trask would put it) for "Pre-Basque Phonology", and that the etymologies "destroyed" by it are not destroyed at all.

Vasco-Caucasic Phonology: An Alternative

It must be emphasized that Vasco-Caucasic (= Macro-Caucasic) phonology is in its infancy. We are in an early stage of research, comparable to the Rask-Bopp-Grimm stages of Indo-European studies. And there are very few workers — as far as I know, Václav Blažek and I are the *only* workers! We do not pretend that we have all the answers just yet, though we are convinced that there are obvious regularities running through our data. Of course, Trask shows no understanding of this preliminary stage, and expects us to be able to explain everything! (Good lord, after two centuries and countless [wo-]man-hours of Indo-European studies, there is still no universal agreement on its phonology!)

What is worse, Trask ignores all the cases that *do* show strong regularity (see the examples above, corresponding to Basque *h*) and focuses instead on the most problematic area: the sibilants and sibilant affricates. Most insidiously of all, he throws in examples that I did not use in my phonology paper, to try to show that there is no regularity. This is tantamount to refuting Grimm's Law by pointing out the exceptions later explained by Verner's Law! We do not deny that there are exceptions, but the general correspondences are clear:

Basque (orthographic)	**s**	= (phonetic)	[ś]	= Caucasic	**c**
	s		[ś]		**č**
	s		[ś]		**š**
	z		[s̄]		**s**
	z		[s̄]		**c'**

Again, as we saw above with Basque *h*, Basque phonemes correspond to five Caucasic phonemes. The phonological evidence is all cross-confirming, and shows clearly how the simple phonemic system of Basque correlates with the more complex Caucasic system.

Vasco-Caucasic phonology is constantly being refined. One of the most exciting recent developments is the discovery of correspondences to the Caucasic lateral affricates (*dˡ tˡ tˡ'*):

Basque *erdera*, *erdel-* 'foreign': Caucasic **dˡōlV* 'guest, neighbor' (see Table 3);

Basque *adar* (< **ardar*) 'horn' : Caucasic **tˡwǐrV* 'horn': Burushaski - *tur*, *-ltur* 'horn';

Basque *medar* (< *merdar*) 'narrow' : Caucasic **ʕVmV̲tˡ'V* 'thin';

Basque *erdi* 'half, middle' : Caucasic **-e̲tˡ'i* 'half, middle' : Burushaski *-alt-* {element inferred in numerals 2, 4, 8)

Basque *urde* 'pig, swine' : Caucasic **wārtˡ'wə* 'pig, swine'.

Basque *lapa* 'burdock' : Caucasic **tˡ'ăpi* 'leaf' : Burushaski *tap* 'leaf', du-*ltapi*-as 'to wither'.

Basque *lur* 'earth' : Caucasic **nĕldˡwǐ* 'earth' (dissim. < **lĕldˡwǐ* ?)

Basque *hur* 'hazelnut' : Caucasic **ħwōtˡ'V* 'nut'

Basque *hil* 'to die, kill' : Caucasic **ʔiwtˡ'V* 'to die, kill'.

These and other examples point to some interesting non-trivial correspondences: both Basque and Burushaski show forms with metathesis of the Caucasic affricates, i.e. *t^{l} > *lt* > *rd*, but these only occur in intervocalic positions, e.g., Basque *adar* 'horn' (from **a-rdar*), and Burushaski *-ltur* 'horn', the form that is used with a possessive prefix. (The dissimilation in Basque *adar*, *medar* is supported by none other than Michelena [1961:339], e.g., Bizkaian *adore* from *ardore*.) In initial and final positions (at least in Basque), the affricates have merged with the usual *l* (see *lapa*, *hil*).

If this is all "chance resemblance", as Trask will probably claim, I challenge him to produce a similar set of non-trivial correspondences between Basque and Hungarian.

Is It Romance?

As we have already seen, Trask's strategy has been to eliminate as many words as possible from the native Basque lexicon, and one of his favorite tactics is to claim that a word is not Basque at all, but a loanword from Latin or "Romance". In a few cases he is correct, but let us look at some of his examples.

In his preliminary remarks on the prehistory of Basque, Trask makes a list of Basque words of purported Latin or Basque origin. Many of these are indeed just what Trask says, but the following have at least tentative parallels in Caucasic:

Basque *hanka* 'leg' : cf. Caucasic **ʔą̄nqV* (Andi *aqu* 'thigh', Tindi *anqu* 'knee-bone'; Archi *aq* 'leg, foot, hind leg (of animal)', etc.

Basque *kokots* 'chin' : cf. Caucasic **q'ăc'ɨ̆* (Lak *q'ac'* 'bite, mouth'; Rutul *q'ac'* 'chin'; Tsakhur *q'ac'* 'chin', etc.).

Basque *potorro* 'vulva', *potro* 'testicle' : cf. Caucasic **pŭt'i*/**bŭt'i* 'genitals' (Chechen, Ingush *bud* 'vulva'; Agul *but'* 'penis', etc.)

Basque *katu* ~ *gatu* 'cat': cf. Caucasic **gā̆tu(jV)*/**gā̆t'u (jV)* 'cat' (Tabasaran, Archi *gatu*; Tsez, Hinukh *k'et'u*; Botlikh *gedu*, etc.)

Basque *azeri* 'fox' (already discussed above).

Basque *abere*, *abel-* 'domestic animal': cf. Caucasic **bül̲V* (Udi *bele* 'cattle'; Chechen *bula* 'aurochs', etc.)

Basque *orga* 'wagon': cf. Caucasic **həlkwē* (Dargi *urkura* 'carriage, cradle'; Lak *hark*^{w} 'carriage wheel axle', etc.)

Basque *lama* 'flame': cf. Caucasic: Lak *lama* 'flame'; Dargi *lami* 'flame'

Basque *ondo* 'bottom' (discussed above).

Of course, not all of these words are important to the Vasco-Caucasic hypothesis. 'Cat' for example is a notorious *Wanderwort*: it is probably a loanword in both Latin and Caucasic, but the Basques may already have had the word before the Romans came. 'Wagon' could be a similar story, at least Latin *organum* does not easily lend itself to the meaning 'wagon'. Basque *hanka* 'leg' has nothing to do with Latin or Romance: it is usually explained as a loanword from Germanic, but what is the mode of transmission? I think the Caucasic parallels look quite good. We may never know for sure about 'flame': both the Latin and the Caucasic parallels are precise, and one might wonder if Russian *plamja* is connected with the Lak and Dargi words.

I would argue more vigorously about Basque *abere*, which Trask glosses as 'domesticated animal', and is traditionally derived from Latin *habere* 'to have', with the idea that 'having, belonging' shifted to 'livestock (possessions)', the reverse of Latin *pecu*, *pecunia*. There are several problems with this; (a) it is unlikely that Basque would borrow the *infinitive* form so exactly (the only example?); (b) the meaning 'belongings, possession' is alien to Basque *abere*: it refers *only* to animals, not only domestic, as shown by the dialectal word *ugabere*, a folk-etymological reshaping of *udagara* 'otter'; (c) the stem alternant *abel-* found in compounds such as *abel-bürü* (Z) 'head of cattle' points to earlier **abele*, not *habere*. As to supposed Romance parallels, their meaning is primarily 'possessions' (e.g., Spanish *haberes* 'property, goods, cash, assets'), not 'animal'.

It is impossible to say yet how all these etymologies will eventually play out. I thought Trask's pronouncement that all words are Latin or Romance, ignoring alternative parallels, should not pass unchallenged.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

Vasco-Caucasic Morphology: Fossilized Class Prefixes

In several of my papers, as Trask notes, I have developed the hypothesis of a Vasco-Caucasic class (gender) system, of which only fossilized traces remain in Basque. I am not the first to suggest this, as Trask also notes, but I believe I have developed the idea farther than anyone else, with much helf from Václav Blažek.

Trask makes a number of irrelevant comments in an attempt to "destroy" the hypothesis. For example, he carps that I do not include *all* anatomical words beginning with *b-* under the *be-/bi-* prefix. If I know that German has a prefix *be-* (not claimed to be connected with Basque *be-*!), is it necessary to derive all word beginning with *be-* as having the prefix, e.g., *Becken*, *Besen*? Of course not, making Trask's argument seem silly. The reason is that for some of these Basque words there is no internal or external evidence that *b-* is a prefix. The *be-/bi-* is only separated when there is internal or external evidence that the remainder of the word is the root. An example of internal evidence is Basque *hatz* and the doublet *be-hatz*, attested as:

Basque *hatz* ~ *atz* 'trace, step', (B) 'finger, toe, thumb', (LN, R) 'paw (of animal)', etc.
Basque *be-hatz* ~ *be-atz* (B) 'toe', (HN, G, LN, Z) 'finger, toe', (HN, G, Z, L, LN) 'thumb', (B) 'hoof, (L) 'nail', etc.
Cf. Caucasic **kwăč'ĕ* (Avar *kʷač'* 'paw'; Hunzib *k'ɔč'u* 'distance between thumb and forefinger', etc.); Burushaski *qaṣ* 'cubit'.

There is no evidence for Trask's claim that *be-hatz* comes from *behe-hatz* 'lower digit'; in fact several dialects use it for 'thumb', which is the *uppermost* digit when the hand is extended or in use. I think these cases are better explained as "stage III articles" (Greenberg 1978), and fit Greenberg's (1987:47) description that "the stage III article . . . appears with some nouns but not with others in a quite sporadic way that differs from language to language." In some cases, the Basque prefixes serve to distinguish words that would otherwise become homonyms: *e-lur* 'snow': *lur* 'earth': *a-hur* 'hollow of hand': *hur* 'hazelnut'.

Trask's objection that there are no living prefixes in Basque is also irrelevant: the argument is that these are *fossilized* prefixes, no longer a part of the living morphological system.

Furthermore, Trask's observation that "three of Bengtson's four classes remarkably include 'body parts' as one of their functions," is not damaging at all. There is no rule that anatomical words all have to belong to the class. For example, in Lak (Laki), an East Caucasic language, 'tongue, nose, mouth', etc. belong to the "*b*-class" (cf. Basque *be-/bi-*), while 'tooth, hair, neck, etc. belong to the "*d*-class" (cf. Basque *ar-*). We know in Indo-European that such words can be masculine, feminine, or neuter.

Since the articles Trask criticizes were written we have discovered an additional prefix, **ar-*, for example:

Basque *ar-han* (LN, Z) 'plum': cf. Caucasic **γōnáV* 'pear' (Hunzib *hĩ*; Avar *geni*, etc.)
Basque *ar-dano* 'wine' (Hualde 1991:85) : cf. Caucasic **ðwənʔɨ* 'wine' (Hunzib *ža* 'wine'; Botlikh *žana* 'winesap'; Lak *ḫan* 'wine', etc.
Basque *ar-ska* 'trough' or 'crib' (see Trask under *aska*) : cf. Caucasic **ć'ä̱q'wV* (Ubykh *čaq'ʷ* 'basin, tureen', etc.); Burushaski *čuq* 'sieve, measure of grain'.
Basque (archaic) *ar-sto* > *asto* 'donkey' : cf. West Caucasic **čʷVdə* 'donkey' Čirikba).

We can provisionally make the following correlations between the Basque fossilized prefixes and Proto-Caucasic class markers (Diakonoff and Starostin 1986:10, 71):

Basque	*u-/o-*	Caucasic I. sg.	**u-*
Basque	*i-/e-*	Caucasic II. sg.	**j-*
Basque	*bi-/be-*	Caucasic III. sg.	**w-*
Basque	*ar-*	Caucasic IV. sg.	**r-*

I must emphasize that this aspect of Vasco-Caucasic morphology is still at a very provisional stage. We can not yet explain everything (as Trask seems to require), but Blažek and I are confident that this is one of the most promising lines of evidence for Vasco-Caucasic and Dene-Caucasic. Trask's criticisms fall apart under examination and do nothing at all to "destroy" this hypothesis.

Conclusions: A New Paradigm for Vasconists

My own evaluation of the 317 comparisons listed by Trask is as follows:

++	probably correct	121
+	problematic, but probably correct, or partially correct	116
-	problematic , and probably incorrect	38
--	probably incorrect	35

(The total is less than 317 because non-Vasco-Caucasic comparisons were left out of the totals, e.g., those with Egyptian, Berber and Kartvelian.) The etymologies I have voluntarily dismissed (- or --) are not necessarily the same ones that Trask claims to destroy. Cognizant that even some of the comparisons I have accepted as (+) or even (++) may eventually be invalidated, we end up with about two out of three of the comparisons being solid.

Two hundred basic comparisons are nothing to dismiss out of hand, the more so when phonetic regularities have been shown (though not yet completely understood), and there are morphological parallels as well. Apart from a limited number of cases where I admit he is correct, the great majority of Trask's objections have been shown to be groundless, growing out of his intransigent zeal to keep Basque isolated at all costs.

Professor Trask, who calls his paper "A Critique from the Basque Side", apparently sees himself as a defender of Basque. The Dene-Caucasic hypothesis seems to present an affront to his view of Basque, an assault on its pristine isolation.

I did not seek this confrontation. Until recently, my efforts were devoted to the presentation of positive evidence. But since Trask's curiously acrimonious attack demands rebuttal, I must state that I believe his isolationist zeal is harmful to the present and future of Basque historical linguistics. His position forces him to reduce Basque to a patchwork of "Romance" elements (many of which cannot be traced to Latin) and native Basque lexemes that for Trask can never be verified by external comparison. Trask's rigid adherence to a logically unsound "Pre-Basque Phonology" makes any external comparison of Basque virtually impossible. As long as he and other vasconists insist on clinging to these precepts, Basque historical linguistics can only lead to stagnation and frustration.

There is a way out. I invite Trask and the other vasconists to *work with us* rather than against us, letting in the fresh air of external comparison. But to do so, they will have to concede that their theoretical constructs ("Pre-Basque Phonology") and widely accepted etymologies must be open to possible revision or even repudiation when viewed in the light of comparison with Caucasic, Burushaski, and other Dene-Caucasic languages. This approach can and will lead to a deeper understanding of all languages concerned, and of their prehistory.

For a brief summary of current Vasco-Caucasic evidence, see the three appended tables. The first consists of some of the clearest Basque-Caucasic parallels, many of which were suggested by Trombetti, Bouda, Lafon, and others. Forms from several representative Caucasic languages are given, as applicable. The Proto-Caucasic forms are those of Nikolaev and Starostin (1991, 1992), which I do not endorse in detail, but consider important first steps. One can see that the meanings are all very basic and usually identical in Basque and Caucasic.

The second table lists some of the Basque-Burushaski comparisons by Hermann Berger (1956, 1959). Some of them coincide with etymologies I independently discovered before I knew of Berger's work. At last report, Berger (personal communication, 1995) continues to work on Burushaski but not on external connections.

The third table details some of the latest work on Vasco-Caucasic phonology. Comparison with Caucasic casts light on the *l/r* alternations, Northeastern Basque *-lh-*, and Basque final *-i*. These phonological patterns are cross-confirming in ways that cannot be attributed to mere chance.

References (in addition to those cited by Trask)

Bengtson, J. D. 1992. "Macro-Caucasian Phonology (Revised Version)". In Vitaly Shevoroshkin, ed., *Nostratic, Dene-Caucasian, Austric and Amerind*, pp. 342-352.

Berger, H. 1956. "Mittelmeerische Kulturpflanzennamen aus dem Burušaski." *Münchener Studien zur Sprachwissenschaft* 9:4-33.

——— 1959. "Burushaski Lehnworter in der Zigeunersprache." *Indo-Iranian Journal* 3/1:17-43.

Diakonoff, I. M. and S. A. Starostin. 1986. *Hurro-Urartian as an Eastern Caucasian Language*. (= *Münchener Studien zur Sprachwissenschaft*, Beiheft 12.) Munich.

Greenberg, J. H. 1978. "How Does a Language Acquire Gender Markers?". In J. H. Greenberg, ed., *Universals of Human Language*. Stanford, Stanford University Press.

——— 1987. *Language in the Americas*. Stanford: Stanford University Press.

Hualde, J. I. 1991. *Basque Phonology*. London/New York, Routledge.

Larrasquet, J. 1939. *Le basque de la Basse-Soule Oriental*. Paris, Klincksieck.

Michelena, L. 1961. *Fonética histórica vasca*. San Sebastián: Diputación de Guipúzcoa.

Morvan, M. 1992. "Remarques au sujet des comparaisons Macro-Caucasiques de John D. Bengtson." *Bulletin de la Société Linguistique de Paris* 87/1:357-365.

Moutard, N. 1975. "Étude phonologique sur les dialectes basques, I." *Fontes Linguae Vasconum* 19:5-42.

Nikolaev, S. L. and S. A. Starostin. 1991. "North Caucasian Roots," in Shevoroshkin 1991, pp. 174-264.

——— 1992. *Caucasian Etymological Dictionary*. Manuscript.

Pedersen, H. 1962. *The Discovery of Language: Linguistic Science in the Nineteenth Century*. Bloomington, Indiana University Press.

Starostin, S. A. 1989. "Nostratic and Sino-Caucasian," in V. Shevoroshkin (ed.), *Explorations in Language Macro-families*, pp. 42-66. Bochum: Brockmeyer.

Tiffou, E. and J. Pesot. 1989. *Contes du Yasin*. (SELAF N° 303.) Paris, Peeters/SELAF.

TABLE 1: BASQUE AND CAUCASIC

General Gloss	Basque	Nakh	Dargi-Lak	Avar-Andi-Tsez	Lezgian	West Cauc.	Proto-Cauc.
DIE	hil	-al-		-ilʔ-[5]		tl'a-[6]	*-iwtl'E
DOG	hor	pħar-	D Xur-	Xʷar-[7]	Xor[8]	ħa[6]	*XHwe̲je
DRY	igar, agor	d-eq'a	L q'arq'-	-eq̄'u-	q'uru	-ʕʷa-[9]	*-i̲GwVr
EAR	be-larri	lar-k'[10]	D liħi			ləmħa[11]	*l̲eHl̲i
EYE	begi, ber-	bʕär-g	D ħuli	ber[1] häre[4]	pul[12] ul[2 15 17 18]	bɮa[13]	*ʔwilʔi
FIRE	su	c'i	D c'a L c'u	c'a[1 3] c'o[4]	c'a[2]	m-ca[11]	*c'aj̲ɨ
FULL	oso	-eza	D -ic'i-	-ec'u-[14]	ac'u[15]	jəz[6]	*ɦoc'V
GO	j-oa-n		L n-an	-əʔa[16] -oʔan-[5]			*HVʔwVn-
HAND	a-hur	ko[10]	L ka	kʷer[1] koro[16]	kul[12 21]		*kwīlʔi
MOUSE	sagu	šat'q'a		sar̄tˡ'u[3]	sok[18]	cəγʷa[6]	*c̲argwɨ̲
NOT	ez	ca		-šu, -š-[3]			*ć'ə / *ʒ́ə
OLD	zahar	šira		-asra-[1]	sur	*śwə	*swɨ̲rHo
STAR	i-zar		D zuri	c̄'wari[24]		*c'wa	*ʒHw̲arhi
STICK	i-dul-ki	täl-g[20]	D t̄al L t̄ala		dali[21]		*dwāl̲ɨ̲
SUN	ekhi	ʕüq'e		q̄'o[1]	iq[21]		*Hwīq'ī
THOU	hi *-ga	ħo	D ħu		γu[18]		*Gu
TOMORROW	bihar ~ bigar			bog[1]	p̄ak̄a bɨga[17]		*bəgə
TONGUE	mihi mintz	mott	D mez L maz	mic̄i[7 14] mic[4]	melz[15]	bza[6]	*mē̲lc'i
WATER	hur			ħor[1] ihir[22]	ür		*ħ̲wirɨ̲
WHAT	ze-r	ste-	D se	se[23]		sa[13]	*s[ā̲]
WHO	no-r			ne-ti[25]	ni na[2]		*ni
WOOD	zur ~ zul			c'ul[1] c'ule[3]		a-c'la[9]	*c'wɨ̄l̲HV
YOU (PL)	zu	šu	L zu		šu[18]	sʷe[6]	*źwV

Notes to Table 1:

DIE: Basque 'die, kill; dead'. Chechen, Karata, Adyge-Kabardian 'die'.

DOG: The Chechen form is the oblique stem; Dargi (dial.) and Godoberi are plural stems.

DRY: Basque forms vary in meaning, e.g., in western Basque *agor* is used in reference to springs or rivers, but *igar* for plants. Another word for dry, *idor*, is probably of a different origin.

EAR: Basque *be-* is probably a fossilized class prefix (cf. Caucasic **w-*). Batsbi *-k'* is a relic of a diminutive suffix: cf. Basque *-ki* in the etymology STICK.

EYE: Basque *begi* [beɣi], *ber-* (attested in a Souletin form *ber-phuru* 'eyebrow') can be traced to **berɣi*, corresponding to **ʔwilʔi*, the Proto-Caucasic form proposed by Nikolaev and Starostin. (*ʔ* is a reconstructed pharyngeal stop.)

FULL: Basque *oso* is properly 'whole, entire, complete'; the Caucasic words all mean 'full'.

HAND: Basque *ahur* (*a-hur* from **a-xur*) is properly 'hollow of the hand'.

MOUSE: Basque 'mouse'; Adyge-Kabardian 'mouse, marten'; Chechen, Andi, Tsakhur 'weasel'; cf. Burushaski *čharge* 'flying squirrel'.

OLD: West Caucasic **śwə* 'year'.

STICK: Basque 'wood block / tronco (pedazo de)'; Ingush 'small stick'; Dargi (dial.) 'pole, post'; Lak 'log, cudgel'; Archi 'long stick'.

SUN: The Basque form is from Souletin; Chechen 'time before dawn'; Avar, Archi 'day'.

THOU: (second person singular pronoun): Basque **-ga* (second person singular intimate agreement suffix, for male, i.e., masculine) is reconstructed by Trask as the predecessor of modern *-k*.

TOMORROW: Avar 'evening, evening meal'; Lezgi, Rutul 'tomorrow', as in Basque. Proto-Caucasic 'morning, evening'.

TONGUE: Basque *mintz* 'voice, word'; all the rest 'tongue'.

WATER: Basque 'water'; Avar, Chamalal, Lezgi 'lake'.

WOOD: Basque 'wood / madera / bois'; Avar 'wood, firewood'; Andi 'big stick'; Abkhaz 'tree'.

YOU (PL): Basque originally second person plural, developed into polite pronoun, parallel with Spanish *usted, vos*, French *vous*.

LANGUAGE IDENTIFICATION

1. Avar
2. Agul
3. Andi
4. Bezhta
5. Karata
6. Adyge-Kabardian (Circassian)
7. Godobere
8. Budukh
9. Abkhaz
10. Batsbi
11. Abaza
12. Udi
13. Ubykh
14. Tindi
15. Tabasaran
16. Hunzib
17. Rutul
18. Tsakhur
19. Kryts
20. Ingush
21. Archi
22. Chamalal
23. Hinukh
24. Akhwakh
25. Tsez

TABLE 2: BASQUE AND BURUSHASKI

General Gloss	Basque	Burushaski	Source
APPLE	sagar [śaɣaɾ̄]	šuɣúri	BLZ 34
BAREFOOT	orthuts ~ ortotz	holtaṣ ~ hultaṣ	BLZ 27
CLOTH	ehun	hunám	BLZ 37
DAY	egun	gun ~ gon	MK 16
DUNG	gorotz	ɣuraṣ	MK 12, BLZ 23
FIGHT	borroka	birga	MK 7
FIRE	su	ši	MK 17
FLY (n.)	euli	ahúlal	MK 16
FOOT	huñ ~ oin	-ɣán	MK 10
HALF/TWO	erdi [eɾ̄ði]	alt-	BLZ 24
HORN	adar	tur, -ltur	BLZ 34
LUNG	hauspo	xurpat ~ xorpet	BLZ 21
MOON	ilaski	halanc	BLZ 36
NETTLE	hausin ~ asun	ɣašu	BLZ 21
NOSE	muthur ~ mustur	-multur	BLZ 33
RAIN	hurolde	haralt	MK 7, BLZ 35
SHOE	oski	ṣoq	BLZ 21
SUN	iguzki	gunc	BLZ 36
THORNBUSH	sasi	ċhaṣ	MK 24
VALLEY-1	ibar	bar	MK 7, BLZ 28
VALLEY-2	haran	har	BLZ 37
WING	hegal	galgi	MK 7
WORM	har	har	MK7

Notes to Table 2: Basque and Burushaski

APPLE: Basque 'apple'; Burushaski 'a kind of pear'; cf. Caucasic: Lezgi *čüX^{w}er*, Tabassaran *ǯeXer*, etc. 'pear'.

BAREFOOT: Basque *ointhuts* secondary by association with *hoin* 'foot' (see table).

CLOTH: Basque 'linen'; Burushaski 'fine cloth (as a gift of honor); bridal ornament'. (Cf. Caucasic: Andi *hon* 'flax'? This is one of the comparisons suggesting a Neolithic dispersal of Vasco-Caucasic. Cf. SHOE).

DAY: Burushaski 'dawn'; cf. *gunc* 'day'.

DUNG: Basque also *k(h)orotz* 'dung (of animals)'; cf. Caucasic: Archi *k'urč'* 'dung of sheep', etc. (The *g-/ɣ-* forms are probably older, the *k-/k'-* forms a more recent assimilation.)

FIRE: Burushaski 'hearth'.

FLY: Burushaski 'moth, butterfly'; the cited form is from Yasin, which equals Hunza *hólalas* id.

FOOT: Burushaski 'heel'; cf. Caucasic **ʔĭnGwV ~ *ʔänGwV* 'heel, ankle'.

HALF/TWO: Basque 'half, middle'; Burushaski *alt-* 'two', also in *walt-* 'four', *altámb-* 'eight', *altar* 'twenty'; cf. Caucasic **-e̲tˡ'i* 'half, middle'.

HORN: Basque probably from **a-rdar* = Burushaski *-ltur* (form used with prefixes); cf. Caucasic **t'wɨrV* 'horn' (Avar *t̄ar*, Chechen *kur*).

LUNG: Basque 'lung' only in the Vizcayan dialect, elsewhere 'bellows'; cf. Caucasic: Andi *Xunsir*, Dargi *Xurhala ~ Xurhari*, etc. (An expressive word with irregularities all around, like Indo-European 'spleen'.)

MOON: Basque is a Labourdin form, derived by Berger from *ilantz-ki*. Cf. SUN. NETTLE: For the correspondence Basque *h* = Burushaski *γ* (voiced uvular fricative), cf. also FOOT. NOSE: Basque 'snout'; Burushaski 'nostril'; cf. Caucasic **mərtˡ'V* 'nose, beak' (Batsbi *marłɨo* 'nose, snot'; Bezhta *motˡ'o* 'beak'). Basque forms are irregular, or contaminated with another root, cf. Chechen, Ingush *muc'ar* 'snout, muzzle, trunk'.

RAIN: Basque 'flood'; Burushaski 'rain, rain cloud'.

SHOE: Burushaski 'sole of a boot'; cf. Caucasic: Tabasaran *šaqʷ ~ šʷaq* 'heel', etc.

SUN: Burushaski 'day'. Berger posits a Basque derivation from **i-guntz-ki*, parallel with MOON, above.

THORNBUSH: Cf. Caucasic: Dargi *čače ~ čanči ~ zanzi*, Tsakhur *zaza*, Chechen *zez*, etc. denoting various prickly plants.

WORM: Basque 'worm, maggot'; Burushaski 'corn worm'; cf. Caucasic **ħabarV* 'worm' (Avar *ħapára*, Tsakhur *ąbra-wuč'*).

References

MK refers to Berger (1956), while BLZ refers to Berger (1959).

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

TABLE 3: VASCO-CAUCASIC PHONOLOGICAL PATTERNS

A. *l* ~ *r* alternations:

Basque *gari* ~ *gal-* 'wheat'	Caucasic **Gōl̲ʔV* 'wheat': Lezgi *q̄ul*, Andi *q'ir*, Akhvakh *q'iru*, etc.
Basque *azeri* ~ *azel-* 'fox'	Caucasic **cEHwōlĕ* 'fox': Tabasaran, Agul *sul*, Udi *šul*, Tsez *ziru*, Tindi *sari*, etc.
Basque *erdera* ~ *erdel-* 'foreign, not Basque'	Caucasic **d̪ōl̲V* 'guest, neighbor': Khwarshi *t̪ar* 'guest', Chechen *lūla-Xō* 'neighbor', Hurrian *lul-a-ḫ·ə* 'foreigner, metoecus', etc.
Basque *abere* ~ *abel-* 'domestic animal'	Caucasic **bül̆V* 'horned animal': Chechen *bula* 'aurochs', Udi *bele* 'cattle', etc.
Basque *hari* ~ *hal-* 'thread'	Caucasic **Xāl̲V* 'sinew; thread': Lezgi *γal* 'thread', Lak *x̄al* 'thread, wire', Tsez *Xero* 'sinew', etc.
Basque *uk(h)are* ~ *uk(h)al-* 'wrist'	Caucasic **q'w[ā̄]l̲ʔV* 'arm; bosom, armpit': Lezgi *q'ula-k* 'in one's bosom', Avar *q̄'ʷal* 'arm', Hunzib *q'eru* 'shoulder', etc.
Basque *zumar* ~ *zumal-* (G) 'elm'	Caucasic **ǯu(H)mal̲V* ~ **ǯu(H)l̲amV*: Budukh *ǯumal* 'cornel (dogwood)', Chechen *ǯʕolam* 'shrub, bush', etc.
Basque *zur* (common) ~ *zul* (B) 'wood'	Caucasic **c'wɬ̄l̲HV*: Avar *c'ul* 'wood, firewood'; Andi *c'ule* 'big stick', Abkhaz *a-c'la* 'tree'.

Each of the eight cases of Basque *l* ~ *r* alternation correspond to a Caucasic parallel with intervocalic lateral. Except for the last example (where Bizkaian *l* corresponds to common Basque *r*), the Basque alternation is morphological (e.g., *gari* has the stem alternant *gal-* in compound words). In Caucasic, as far as I know, the alternation is only geographical (dialectal). Nikolaev and Starostin reconstruct two lateral resonants, Proto-Caucasic **l* and **l̲*, with slightly different reflexes in the dialects (e.g., **-l-* becomes *-r-* in Avar-Andian and Tsezian), but we see no evidence for the distinction in Basque.

There is also the Caucasic unvoiced lateral fricative: **ɬ*, which corresponds to the intervocalic "French" Basque (LN, L, Z) *-lh-*:

Basque (Z) *elhür* 'snow'	Caucasic **ɬwiu̯V* 'snow': Chechen *lo*, etc.
Basque (Z) *elhe* 'to speak'	Caucasic **ʔoɬV* 'to speak': Chechen *ala-r*, etc.
Basque (Z) *ilhinti* 'firebrand'	Caucasic **ɬwindV* 'firewood': Andi *ɬudi*, Chamalal *ɬunni*, etc.
Basque (Z) *olho* 'oats'	Caucasic **ɬwɨʔwV*: Lezgi *fu* 'bread', Rutul *xɨw* 'bread', Kabardian *xʷə* 'millet', etc.

B. Basque final *-i* = Avar-Andian final *-i*:

Comparison of Basque with Caucasic has revealed some precise correspondences of final vowels, particularly *-i*. Some of the most striking cases involve the Avar-Andian languages, e.g.:

Basque *mihi* 'tongue'	Avar-Andian **miᾱ'i* 'tongue': Andi, Akhvakh, Botlikh *miᾱ'i*, Tindi, Godoberi *miᾱi*, Karata *maᾱ'i*;
Basque *behi* 'cow'	Avar-Andian **biᾱ'ʷi(r)* 'cattle': Avar *bóᾱ'i*, Andi *buᾱ'ir*, Akhvakh *boᾱ'iri*, Botlikh *buᾱ'i*, Godoberi *purᾱi*;
Basque *azegari* ~ *azeri* 'fox'	Avar-Andian **sori* 'fox': Akhvakh *šari*, Tindi, Botlikh, Godoberi *sari*, Karata *sare*;

Basque *erdi* 'half, middle'	Avar-Andian *-*oɬ̄'i* ~ *-*aɬ̄'i* 'half, middle': Akhvakh *b-aɬ̄'i-s* 'half', Karata *b-aɬ̄'i* 'between', Tindi *b-aɬ̄'i* 'amidst' (*b-* is a gender prefix);
Basque *ilhinti* (Z) 'firebrand, ember'	Avar-Andian **ɬundi* 'firewood': Andi, Godoberi, Tindi *ɬudi*, Karata *ɬure*, Chamalal *ɬunni*, Botlikh *hudi*.

These five comparisons are not only phonetically and semantically plausible, but most remarkably, the final vowels in Basque and in Avar-Andian (a Caucasic family) are *identical.*

Some linguists have told me that they have difficulty accepting the equations *mihi* = *mic̄'i*, *behi* = *boc̄'i* because of the dissimilarity of the Basque aspirate *h* (still audible in "French" Basque) and the Caucasic glottalized affricate *c'* [ʦ']. Some time ago (Bengtson 1994b:33), I suggested a development of the type:

*milc̣i	>	*miṣṣi	>	*mixi	>	mihi 'tongue'
*berc̣i	>	*beṣṣi	>	*bexi	>	behi 'cow'

The development of *sh*ibilant *š* or retroflex *ṣ* to a velar fricative (*x*, x^w) is widely attested (e.g., Spanish, dialectal Swedish, Slavic). Only recently have I learned that the **x* hypothesized is in fact preserved in the western Low Navarre (Basse-Navarre) dialect of Basque, in the forms corresponding to standard Basque *mihia* 'the tongue' and *behia* 'the cow', which are transcribed by Moutard (1975:27) as:

[mih(ç)ja] 'la langue'
[beh(ç)ja] 'la vache'

I interpret this [ç] (which is of course the fronted variant of [x]) as a localized conditioned retention of an archaic sound, particularly since Moutard also notes the presence of [x] in [a(x)hwa] 'the mouth' (orthographic *ahoa*), verifying the validity of the reconstruction **a-x*^{w}*o* 'mouth (Bengtson 1994b:34).

These phonological patterns (and others could be added) refute Trask's contention that Vasco-Caucasic and Dene-Caucasic comparisons are all merely "random resemblances". Random resemblances do not behave in this patterned way. Note that some of the above comparisons bear witness to two or more different patterns, e.g., Basque *azeri* ~ *azel-* 'fox' has both the *r* ~ *l* correspondence to Caucasic **l*, and the correspondence of final vowel (Avar-Andian **sori* 'fox'). Zuberoan Basque *lihinti* shows at least two regular correspondences with Avar-Andian (*lh* = *ɬ* and *-i* = *-i*).

Note: Caucasic data come from Diakonoff and Starostin (1986) and Nikolaev and Starostin (1991 and 1992).

Transcription Guide:

ã, ẽ, etc.	nasalized vowels
a̰, ḛ, etc.	pharyngealized vowels
a̲, e̲, etc.	prosodic condition of vowels (tense voice?, see Starostin 1989)
c, c', ʒ	[t͡s], [t͡s'], [d͡z], dentalveolar affricates
c̄	[t͡s̄], Basque dorso-alveolar affricate (orthographic <tz>)
ć	[t͡ś], Basque apico-alveolar affricate (orthographic <ts>)
č	[t͡š], Basque prepalatal affricate (orthographic <tx>)
G = g	voiced uvular stop
γ	(Basque) voiced velar fricative; (Caucasic and Burushaski) voiced uvular fricative
h	glottal fricative (audible in "French" Basque; silent in "Spanish" Basque)
ħ = ḥ	voicless pharyngal fricative
ʔ = '	glottal stop
ʔ̣	pharyngal stop
ʕ	voiced pharyngal fricative
H	"laryngeal" of undetermined quality
i̱	high mid vowel
l̲	lateral sonant (possible velar or "dark" *l*)
ɬ	voiceless lateral fricative (<hl>)
dˡ = h	voiced lateral affricate
tˡ, tˡ' = ħ, ħ'	voiceless and glottalized lateral affricates
p̄, t̄, k̄, s̄ = p:, t:, k:, s:	fortis ("emphatic") consonants
p', t', k', etc.	glottalized consonants
q, q'	voiceless and glottalized uvular stops
r̄	Basque trilled resonant (orthographic <rr>)
s̄	Basque dorso-alveolar fricative (orthographic <z>)
ś	Basque apico-alveolar fricative (orthographic <s>)
š	prepalatal (postalveolar) fricative (Basque orthographic <x>)
ṣ	retroflex fricative (Burushaski)
x	(Basque) voiceless velar fricative; (Burushaski) voiceless uvualr fricative
ü = [y]	"French" Basque and Chechen fronted *u*
V	reconstructed vowel of undetermined quality

Definitions:

Burushic	a small language family in the far northwest of greater India, now represented by the Hunza-Nagir and Yasin dialects of Burushaski. A Burushic substratum can also be traced in certain Indo-European languages, e.g., Khowar and Shina.
Caucasic	native languages of the Caucasus region, exclusive of Indo-European, Altaic (Turkic, Mongolic), Semitic, and Kartvelian languages. They can be positively described as (North) Caucasic, i.e., Abkhazo-Adygan and Nakho-Dagestanian, as well as the extinct Hurrian, Urartean, and Hattic (Diakonoff and Starostin 1986).
Dene-Caucasic	(also called Dene-Caucasian, Sino-Caucasian) a macrophylum consisting of the Vasco-Caucasic (see below), Sino-Tibetan, Yeniseian, and Na-Dene phyla or families. The possible inclusion of Sumerian and Kusunda is less certain.

Vasco-Caucasic (also called Macro-Caucasic, Macro-Caucasian) a phylum level subdivision of Dene-Caucasic, consisting mainly of Basque, Caucasic, and Burushic, as well as extinct languages such as Aquitanian, and possibly Iberian.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

TOWARDS THE POSITION OF BASQUE: A REPLY TO TRASK'S CRITIQUE OF THE DENE-CAUCASIAN HYPOTHESIS

Václav Blažek
Ryneček 148, 26101
Příbram, Czech Republic.

[Editorial preface: Very small changes in format and very small editorial changes have been made in Professor Blažek's text to fit our format. (Václav recently joined the faculty at the University of Brno.)]

I am not sure how to be sufficiently objective in judging Trask's crushing critique, since the criticized John Bengtson and Vjačeslav Čirikba are good friends of mine. On the other hand, I prefer a rational criticism more than an uncritical enthusiasm in agreement with the motto: "If you want to convince your critics, you must be the strictest critic of yourself." (Vladimir Skalička)

From this point of view, I must agree with some of Trask's objections, while his other estimations seem to be too categorical or aprioristic. So, Trask quotes the discrepancy between the length of the word/root in Basque (... never consists of a single consonant ...) versus North Caucasian (... often no more than a single consonant ...). But there are more monoradical words in Basque, accepted even by Trask: pronouns *ni*, *hi*, *gu*, *zu*, further (for example) *atso* 'old woman', *atzo* 'yesterday', *idi* 'ox, *or* 'dog', etc. On the other hand, the monoradicalism is typical (not obligatory!) only for North-West Caucasian (= Abkhaso-Adygean) languages, while for North-East Caucasian (Nakh-Dagestanian) languages longer word/roots are characteristic. The contrast between 16 pre-Basque consonants and 101 proto-Lezgian consonants is almost comparable with, for example, 8 Hawaian or 9 Tahitian consonants inherited from 46 Proto-Austronesian consonants reconstructed by Dyen (1971).

In Trask's relatively detailed evaluation of preceding attempts to find the external relative of Basque, I miss, e.g., the following titles (referring to various branches of the hypothetical Dene-Caucasian macrophylum): Holmer 1947 and 1953 (a postulation of the Paleo-Eurasian chain of residual languages, consisting of Basque, Caucasian, Burushaski, Yeniseian, Chukchi-Kamchatkan; his arguments are based on typological and morphological similarities); Berger 1956 and 1959 (Basque-Burushaski lexical comparisons); Tailleur 1958 (Basque, Caucasian and Yeniseian lexical and morphological comparisons); Urreiztieta-Rivera 1980 (an overview of Basque-Caucasian comparisons); Furnée 1982 and 1989 (Basque is compared with Burushaski, Kartvelian, and a pre-Greek substratum).

The criticized Čirikba and Bengtson rather differ from their predecessors. They separate Kartvelian from North Caucasian, using frequently the North Caucasian reconstructions of Starostin and Nikolaev, instead of isolated examples from various languages. They try to find the regular phonetic correspondences among presented cognates, fully in agreement with a generally accepted methodology.

Trask is certainly right in excluding borrowings from the play. When he can offer an evident source, all is okay. Also the examples from Iberian or Aquitanian onomastics are valuable for a preservation of a more archaic shape of related Basque forms. But his negative argumentation is frequently based only on his own postulates of Basque historical phonology. It could be useful to repeat the chronology of the most important phonetic changes, mapped thanks to Latin/Romance borrowings. Guitter (1989:800) presents the following sequence of changes:

A.D	200	300		400-1000		1100
	*p- > b-					
	*t- > d-	*mp > mb	*n > Ø	*nt > nd	*l > r	*ll > l
	*k- > g-			*nd > n	*nn > n	
		*mb > m				

Trask, following Michelena and Anderson, excludes the originality of *m* in Basque, supposing its creation in the cluster **nb* / **mb* > *m*, attested already in some Iberian texts (Ascoli Bronze — see Anderson 1988:14) or after *u* or thanks to a nasal assimilation besides Celtic and Latin/Romance borrowings or "expressive" words. But from the point of view of a phonological typology, a language without *m* is rather strange. There is no reason to interpret a scarcity of the sign *m* in Iberian orthography (Cf Anderson 1988:31, footnote 4) as reflecting a total absence of the sound *m* in Iberian and pre-Basque. For example, the Egyptian script did not have any sign *l*, substituting other signs for it (*n* / *ꜣ* / *j*)[1], but there is a full agreement on the presence of the sound *l* here, attested only in Demotic and Coptic orthography. Similarly, in cuneiform spelling of Elamitic and frequently of Akkadian, *m* was used also for *w*.

It is not possible to comment on all 317 etymologies discussed by Trask. The following notes represent only a selection.

5. *adin* 'age', compared already by Berger (1956:16) with Burushaski *den* 'year'. This can be supported by Iberian *adin -/-adin* in (*Adin-gibas*, *Balci-adin*). Aquitanian *Dann-adinnis* (Anderson, 1988:111-112), functionally parallel to Celtic compounds of the type of Gaulish *deae Setloceniae*, Old Brythonic *Vendesetli*, Middle Welsh *Gwynhoedl* with the second component corresponding to Latin *saeculum* (AT, ASJU 22/1:297). I believe this comparison belongs to the most convincing.

10. Michelena's and Trask's reconstruction Basque *ahizpa* 'sister (of a woman)' < **anizpa*, where **an-* can be connected with *anai* 'brother', *ahaide* < **anaide* (?) and *-pa* is perhaps a variant of *-ba* terminating more kinship terms (e.g., *alaba* 'daughter', *arreba* 'sister (of a man), *iloba* 'niece, nephew, grandson/granddaughter', *izeba* 'aunt', *neba* 'brother (of a woman)', *osaba* 'uncle'), allows the separation of the root **-iz-* (cf *izeba* and *izeko* 'aunt') compatible with Dene-Caucasic counterparts.

11. Basque *aho* 'mouth' with variants *a(h)u*, *a(g)o*, and *abo* reflects rather a prototype **aH^{w} o* or the like, rather than 'zero-radical' **ao* (!). Accepting the first reconstruction, the comparison with Yeniseian **Xowe* and Sino-Tibetan **khoow* 'mouth' looks quite plausible. The North Caucasic example **k'wĕl̲ʔV* idem and Sumerian *gu$_7$ /kú* 'to eat' must be rejected (LDC #7).

12. The reconstruction *ahüñe* 'kid' < **anu-* has an alternative in the dialect form (Salacenco) *aguña* (AT, ASJU 23/2[1989]:496), admitting derivation from Germanic **auna-* 'lamb; to wean' (Naert, *Orbis* 12 [1963]:199). Accepting it, the hiatus *-h-* in intervocalic position can reflect not only **-n-*, but also perhaps **-w-* and/or some (labialized ?) laryngeal (cf. #11).

15. Hurrian is an extinct language from the 2nd millennium BC. Together with a related Urartean from the 1st millennium BC, they probably represent a branch of North Caucasian. Their closeness, suggested by Diakonoff and Starostin 1986, must be taken with certain reservations. (Cf. Smeets, *Bibliotheca Orientalis* 46/3-4 [1989]:259-279). But a genetic relationship remains indisputable. It means the comparison of Basque *alaba* 'daughter' and Hurrian *ela* 'sister' is quite legitimate.

20. Trask's reconstruction *amentz* and *amentx* 'gall oak' < **abentz* is fully compatible with North Caucasic **mĕndžV* 'oak', regardless of the priority of either *b* or *m*. Accepting Trask's point of view, we can replace the equation [Basque *m* = North Caucasic **m*] with [pre-Basque **b* = North Caucasic **m*].

32. In LDC #92, we add Hurrian *taṣkar-* 'boxwood' is cognate with (=) Sino-Tibetan: Tibetan *stag-pa* 'birch-tree' = Yeniseian: Pumpokol *teksul-či* 'stem', etc.

[1] Editor's note: This example of 'j' is in the Indo-European tradition. It is [y] in other traditions. Written otherwise, this group of three might be /n/, /ʔ/, /y/.

42. The Basque morph *-ba* forming more kinship terms (see #10) was also analytically separated by Zycar' (1988:8-13), who supposed two homonymical roots, viz., (1) 'son/daughter' and (2) 'brother/sister'.

46. Basque *bargo* resembles Germanic forms, particularly: Old Icelandic *borgr* 'verres', Old High German *baruc*, Old English *bearg* 'barrow', etc. Or Portuguese *bácoro*, Gallician *bacro* 'pigling' (Schuchart, sèe AT, ASJU 24/3 [1990]:850). The cultural character of the word supports a conclusion of borrowing rather than an old heritage. (Editor's note: Schuchardt's 'pigling' is Germanized English. Try 'piglet' or 'shoat' or 'lil piggy'.)

66. Basque *biga* 'two' (with variants *bida/bia*) probably includes the same suffix *-ga* as the other cardinals, i.e., *baga* 'one' < **bat-ga*, *biga* 'two', *higa* 'three' < **hiru-ga*, *laga* 'four' < **laur-ga*, etc. Bengtson's speculation implicates a development *$G^{w}i$ + -ga > *gi-ga / *bi-ga > biga.*

70. The Iberian *bioS* (Anderson, 1988:124) gives more hope for the comparison with Burushaski *-As* 'heart, mind'. Of course, it only works if we accept Uhlenbeck's hypothesis of the body part prefix *b-*.

73. Perhaps a more plausible cognate to Basque *bizar* 'beard' can be East Caucasic[2] **booldzʸV / *bildzʸV* 'beard' > Bežtin *bišal-*, Agul *mužur*, etc. (Nikolaev and Starostin, 1992:55-56).

75. The identiication of the "body part prefix" *bi-* in Basque *bizkar* implies naturally a syncope caused by the prefix. The syncope is not unknown in Basque. Cf. *erdi* 'half' versus Iberian *eredi*, etc. This form can be supplemented to #92 (Anderson, 1988:122).

127. In LDC #120 Yeniseian **čɨʔs* 'stone' is compared with Burushaski *čiš'* 'mountain, hill' and Na-Dene: Eyak *čiiš* 'gravel', etc., while Basque *haitz* 'rock, crag' is left (i.e., taken out of that etymology: Editor).

131. The reconstruction of pre-Basque **karri* 'stone' can be supported by old place names, to wit: *Bolquera* (Cerdaña), *Bolcharia* (876), *Bolcaria* (937), which equal **bulu karria* 'head and stone' (Guitter, 1989:800). The same root is wide-spread in Romance languages as a substratal word (AT, ASJU, 23/1 [1989]:183-4). Contemporarily, it is an independent proof that *h* can reflect an older velar stop, as Trask admits here, but rejects in the case of #145 *hi* 'thou', although it is supported by the verbal marker of the 2nd person *-k*.

149. Anderson (1988:124) compares Basque *hitz* 'word' with Iberian *-iS* in *iSbinai*.

178. Bynon (1984:264-265) admits the influence of Arabic *ism* 'name' in some Berber languages, e.g., Beni Snus *līsəm* (where the /li-/ is the borrowed Arabic article), while in Zenaga, for example, there exist the original Berber *éšəm* 'name' plus the borrowed *élisəm* in the more specialized meaning of 'name in a book'.

252. Basque *sagar* 'apple' is probably of Latin origin. Cf. *malum sacrum* (Griera, 1960 — see Markey in *JIES* 16 [1968]:63-4, footnote 5).

308. Geographically closest parallels to the Basque *zilar* 'silver' (note the variants *zirar*, *zidar*, *zilhar*) are known directly from the Iberian peninsula. See the legend *Salir* occurring on Iberian coins and Celtiberian (Botorrita) *Silapur* (accusative singular) 'silver' (Eska, 1989:97, 188, footnote 50; Meid, 1993:113-114). It is evident that the

[2] This term depends on the separation of Kartvelian from North Caucasic; otherwise it would be — and traditionally was — called Northeast Caucasic. The editor (HF), as is his habit, adheres to the older tradition of calling the two phyla *Caucasic* rather than Caucasian, since the latter is polysemic.

latter form has its closest cognate in Germanic **silubra-* 'silver', while pre-Slavic **sirabro-*, Old Prussian **siraplis*, (accusative) *siraplan*, and East Baltic **sidabr-/*sudrab-* are less close. (W. Smoczyn'ski even judges that the Baltic forms are borrowed from Germanic [personal communication]). These forms are etymologically obscure, including perhaps related Lydian (?) *Sibros· argúreos*. Recently Trubačev (*Voprosy jazykoznanija* 1985/5:5) has presented a new etymology based on a hypothetical Indo-Aryan compound **śubhri-apa-* 'light water'.

Usually quoted "Berber" *azerf* 'silver' as a source or a cognate of Basque *zilar*, as early as Gabelentz (1894:118-119), has a quite different origin. The real forms are, as follows: (South Tuareg) Ahaggaren, Ghat, Ayr, Iullemidden, Taitog *az'rĕf*, Ghat also (à la Basset) *az'ruf* (leading to Hausa *azurufa/azurfa*), Kel Wi *azer*; (West) Zenaga *azerfi* and *azruf*; (North) Beni Menacer, K'sur, Harawa, Warsenis, Shawiya *azerf*, etc. 'silver'; (South) Ghat *az'arif*; (East) *az'arif*; (North) Shilha *az'arif*, Wargla *zarif*, Zayan *azalif*, etc. 'alum'. These are all of Semitic origin via Punic, although in Phoenician this root is attested only in *mṣrp* 'smelter': cf. Hebrew √*ṣrp* 'to smelt, refine', Old Aramaic *ṣryp* 'purified', *ṣrp* 'goldsmith', Ugaritic *ảbn ṣrp* 'stones of *ṣrp*' used for adorning jewelry and garments, Sabean *ṣrf* 'silver', Akkadian *ṣarpu* 'purified silver', *ṣarrapu* 'goldsmith', etc. See Basset (1895:64-66), Vycichl (Aegyptus 38 [1958]:149-150), Gordon (1965:#2199), Aistleitner (1965:1 #2360), and Klein (1987:557).

It is evident that this Berber / Semitic cultural word can be a source neither of Basque nor Iberian nor of North Indo-European. But in the case of Iberian/Basque, there is another Semitic term, namely, Ugaritic *ṣrr* 'gold' and Akkadian *ṣarīru* 'rotglanzende Goldlegierung', which could serve as a denotation of 'silver' on the Iberian peninsula, if exported to Iberians by Phoenicians.

Independently of Trask's strict criticism, I am not so optimistic concerning a number of hopeful cognates connecting Basque with other Dene-Caucasian languages as my friends John Bengtson and Vjačeslav Čirikba are. I even presented above some of my specific objections (#46, #252). On the other hand, I am convinced that the following comparisons are at least promising: 4, 5 ,7, 10, 14 (the isolated Gascon *lagagno* can be of substratal origin), 15, 20, 24, 32-35, 42, [45], 49, 51, 54, 55, 57, 64, 65, 70, 73 (see above), 78, 83 (Burushaski *goon/gUn* 'dawn', etc.), 84, 87, 93, 96, 101, 103, 120, 121-123, 126, 131, 134, 142, 145, 147-149, 154-156, 161, 172, 173, 175, 176, 178 (only Burushaski *sen-As* 'to name'), 185, 187, 189, 190, 191, 198, 199, 206, 210, 217 (Basque *moto* 'head scarf'), 220 (Basque *huñ*), 221 (Gallo-Romance **mūla* has no Indo-European etymology), 224, 227-230, 236, 240, 251, 253, 254 (accepting Trask's reconstruction of pre-Basque **saCats*, where *C* can be a velar too), 260-263 (Burushaski *ši* 'fire-place, etc.'), 272 (Lallwort?), 276, 279, 284, 288, 290-293, 302, 312, 316, 317.

The question of noun-class prefixes is probably crucial for a proof of a genetic relationship of Dene-Caucasian. Till the present time, the class prefixes are alive in Burushaski, Kusunda, and some Nakh-Dagestanian languages, reconstructable for Proto-Yeniseian, while in Abkhaso-Adygean, Sino-Tibetan, Na-Dene, and Basque there are only certain traces of fossilized "class markers". The semantic ambiguity (from the point of view of modern civilization) is typical even for languages with a developed system of class markers, for example, Bantu. But a detailed discussion of it calls for much more space. I expect it in the reply of John Bengtson.[3]

One of the most important criteria for a genetic relationship is a common system of personal pronouns. Let us compare the first and second persons in some Dene-Caucasic branches:

	1st sg.	2nd sg.	1st pl.	2nd pl.
Basque: ind.	*ni*	*hi*	*gu*	*zu(e)*
verb.	*-t (-da-)*	*-k* (m.) (**-ga-*?) *-n* (f.)	*-gu*	*-zu(e)*

[3] Editor's note: The last two sentences apparently mean that class markers lead to semantic ambiguity and/or are a sign of less developed / primitive languages. While that seems extraneous to our taxonomic discussion, such a topic will exercise anthropologists!

		1st sg.	2nd sg.	1st pl.	2nd pl.
Caucasic I	a.	**zoo(-n)*	**uo(-n)*	i. **ǰü*	**ẑwA*
	e.	**ʔez(V)*		e. **ži*	
	II	**nǐ*	**Gu/*ʔVG*		
Hattic		*fa-*	*we- / u-*	*ni-*	
Hurrian	I	*iz̲a-s̲*	*we-*		
	II	*-iff-*			
Burushaski	I	*ža*	*un*	*mi*	*ma*
	II	*a-*	*gu- / go-*	*mi-*	*ma-*
Yeniseian	I	**ʔadz-*	**ʔaw/*ʔu*	**ʔadz-ang*	**ʔaw/k-Vng*
	II	**ʔab-/*b-*	**ʔuK-/*Ku-*		
Sumerian	(Emegir)	*g̃á*	*za / zé*		
	poss.	$g̃u_{10}$	*-zu*	*-me*	*-zu(e)ne(ne)*
	(Emesal)	*me*	*ze*		
Kusunda		*či/ki*	*nu*	*čoo/toʔi*	*noki/noʔi*
Sino-Tibetan	I	**jøʔ* (Karen)	**na[H/n(g)]*		
	II	**ngaa/*ngay* < **ni* (Naga) + **Kay* (Kuki, Dhimal)	**K^{w}Vj*		
Na-Dene:					
Haida	ind.	*ɬ*	*da*	*t!ala'ñ*	*dala'ñ*
	obj.	*dii*	*dañ*	*il!*	*dala'ñ*
	poss.	*naa'-ga*	*da'ñ*	*ii'laa-ga*	*dala'ñ-ga*
Tlingit	ind.	*xa*	*waeʔ*	*u-haa'n*	*/yi-h/waa'n*
	poss.	*x(a)*	*i*	*tu*	*/yii*
Eyak		*x^{w}-*	*ʔi*	*daa-*	*khuuinkhan*
Athabascan		**ši*	**ni*	**nu-hani* **-iiʔd*	**(n)u-hi*

Abbreviations used: a. = absolutive, e. = ergative, ind. = independent, obj. = objective, poss. = possessive, verb. = verbal. [The last means an affix to verb forms: Editor].

A basis for this synopsis is borrowed from the unpublished manuscript of J. Bengtson, "Notes on Sino-Caucasian", 1990. Cf. also Bengtson, 1994.

In spite of a vast diversity of forms, it is possible to find at least traces of a common system. The following quasi-reconstructions are naturally only impressionistic.

		ergative	absolutive	possessive
Singular	1st	**?ad(z)i*	**d(z)u*	**?awi*
	2nd	**?awu*	**wi*	**?uGi*
Plural	1st		**u*	
			**Zui*	

The forms beginning in *n-* can be perhaps explained as reinforced by a prefix **n(i)-*, reflected, for example, in Sumerian *ni* 'self'. In the case of first person, they could be projected into a pronoun of the first person plural, exclusive (e.g., me and you, as opposed to you + me + others). Although the preceding thoughts are only speculations, it is evident that the Basque pronominal system agrees fairly well with the hypothetical proto-system.

Abbreviations for references

ASJU = *Annuario del Seminario de filologia vasca "Julio de Urquijo"*.
AT = Agud, M. and A. Tovar. 1988f: *Diccionario etimológio vasco. ASJU* 22/1f
JIES = *Journal of Indo-European Studies*.
LDC = *Lexica Dene-Caucasica*. See Blažek below.

References

Aistleitner, J. 1955. *Wörterbuch der ugaritischen Sprache*. 2nd edition. Berlin, Akademie-Verlag.

Anderson, J.M. 1988. *Ancient Languages of the Hispanic Peninsula*. Lanham, University Press of America.

Basset, R. 1895. "Les noms des métaux et des couleurs". *Mémoires de la Société de linguistique de Paris* 9:55-91.

Bengtson, J.D. 1994. "Edward Sapir and the 'Sino-Dene' Hypothesis". *Anthropological Science* (Japan) 102/3:207-230.

Berger, H. 1956. "Mittelmeerische Kulturpflanzen aus dem Burušaski". *Münchener Studien zur Sprachwissenschaft* 9: 4-33.

———— 1959. "Die Burušaski-Lehnwörter in der Zigeunersprache." (Unknown journal 3: 17-43)

Blažek, V. and J.D.Bengtson. 1995. "Lexica Dene-Caucasica". *Central Asiatic Journal*. Forthcoming.

Bynon, J. 1984. "Berber and Chadic: The Lexical Evidence", in: *Current Progress in Afroasiatic Linguistics. Papers of the 3rd International Hamito-Semitic Congress*, ed. by J. Bynon, pp. 241-290. Amsterdam-Philadelphia, Benjamins.

Diakonoff, I. M. and S. Starostin. 1986. *Hurro-Urartian as an Eastern Caucasian Language*. München, Kitzinger.

Dyen, I. 1971. "The Austronesian Languages and Proto-Austronesian", in T. Sebeok, ed., *Current Trends in Linguistics*, vol.8: *Linguistics in Oceania*, pp. 5-54. The Hague, Mouton.

Eska, J.F. 1989. "Towards an interpretation of the Hispano-Celtic Inscription of Botorrita". *IBS* 29. Innsbruck.

Furnée, E. 1982. "Lexikalische Beziehungen zwischen Baskisch, Burušaski, Kartvelisch, und Vorgriechisch". *Georgica* 5:27-32.

———— 1989. *Urbaskisch und Urkartwelisch. Indizien für Verwandtschaft*. Leiden, The Hakuchi Press.

Gabelentz, G. von der. 1894. *Verwandtschaft des Baskischen mit den Berbersprachen Nord-Afrikas*. Braunschweig, Satler.

Gordon, C. H. 1965. *Ugaritic Textbook. Glossary and Indices*. Pontificium Institutum Biblicum.

Guitter, H. 1989. "Elementos de cronología fonética del vascuence". *ASJU* 23/3:797-800.

Holmer, N. 1947. "Ibero-Caucasian as a Linguistic Type". *Studia Linguistica* 1:11-44.

———— 1953. "Further Traces of Paleo-Eurasian". *International Anthropological and Linguistic Review* 1:160-178.

Klein, E. 1987. *A comprehensive Etymological Dictionary of the Hebrew Language*. New York, NY, and London, Macmillan.

Löpelmann, M. 1968. *Etymologisches Wörterbuch der Baskischen Sprache*. Berlin, de Guyter.

Meid, W. 1993. "Die erste Botorrita-Inschrift". *IBS* 76. Innsbruck..

Nikolaev, S. and S. Starostin. 1992. *A Caucasian Etymological Dictionary*. Unpublished manuscript.

Tailleur, O.G. 1958. "Un ilot basco-caucasien en Sibérie". *Orbis* 7:415-427.

Urreiztieta-Rivera, I. 1980. *Basque and Caucasian: A Survey of the Methods Used in Establishing Ancient Genetic Affiliations*. Doctoral dissertation, The University of Arizona.

Zycar', J. V. 1988. *Rekonstrukcija v oblasti baskskogo jazyka*. Tbilisi, Mecnierba.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

COMMENT ON TRASK'S CRITIQUE OF "DENE-CAUCASIAN-BASQUE RELATIONSHIP"

Eric de Grolier, First Secretary
Language Origins Society
3, rue Saint-Pierre
45320 Courtenay, FRANCE.

[Editor: Very small changes in format and very slight editorial changes have been made in Professor de Grolier's text to fit our format.]

Professor Trask's scathing critique is a nice specimen of the *Schadenfreude* which is apparently the common mood of traditionally minded historical linguists when they come across attempts to transgress the limits of their own specialized knowledge and current dogmas. It is full of expressions of contempt: many etymologies examined are declared "destroyed", segmentations operated by the unfortunate author(s) criticized are "outrageous" (#58, 100) or "preposterous" (#211), proposed relationships are "absurd" (#212-13) show "monumental confusion" (#83 — this confusion arising from the fact that Professor Trask has combined in this item different comparisons made at different dates by each of the three authors criticized) and, finally, the whole enterprise is "utterly pointless" (p. 104; the same term being used for # 270). Such deprecatory vocabulary may be considered as a proof of the immaturity of social sciences in general and (historical) linguistics in particular, as it has been practically now abandoned in natural sciences, after having been employed there against Darwin in the 19th century, and still against Freud or Wegener in the beginning of our own (to name some scholars of much higher status than Bengtson, Blažek, or Ruhlen).

A detailed critique of this critique would necessarily require more space than the critique itself: it took me about three weeks for scanning it, still rather superficially, and an equal time would probably be necessary for being reasonably complete.

It is easy to subscribe to many of Professor Trask's specific objections to the relationships of Basque words or grammatical features proposed by Čirikba, Bengtson, Blažek, or Ruhlen with (North) Caucasian and/or Sino-Tibetan plus Yenisseian and Na-Dene plus Burushaski and Sumerian — the limits of these comparisons differing according to the particular extension of the supposed phylum or superphylum hypothesized by each author.

One must admit that the "three musketeers" of the younger generation of subscribers to Trombetti's monogenetic theory of *Homo sapiens sapiens* languages — Ruhlen, Bengtson, and Blažek — are prone to make many mistakes in the details of their comparisons. I could propose a long list of errors or misinterpretations on their part concerning Burushaski or Sumerian — the two languages which I have studied in some detail — paralleling that of Professor Trask for Basque.

For Burushaski, it is very regrettable that they apparently used only Lorimer's works, and not the more recent and much more precise works of Professor Herrmann Berger (1974 for Yasin; in print for Hunza/Nager), so that they have been induced to admit as native Burushaski forms a lot of loans from Shina, Khowar, and/or Urdu.

For Sumerian, they did not pay sufficient attention to the existence of different "strata" in its vocabulary; it is dangerous to admit for 'long range' comparisons forms which are only attested in relatively recent (post Neo-Sumerian) sources. Moreover, they do not take into consideration the frequent alternations of forms, which in rather numerous cases would involve a reconstruction different from the form(s) retained by them.

Judging the value of Bengtson's "Macro-Caucasic" is particularly difficult, as the most important piece in his construction is (North) Caucasian, and up to now the ordinary reader has no access to the infrastructure supporting the Nikolaev-Starostin reconstruction which he uses (for my part, I have only the bare list, without details, "edited and arranged" by A. Eulenberg 1990, kindly supplied by Professor Shevoroshkin).

Professor Trask appears particularly infuriated by the fact that the authors criticized treat certain segments at the beginning of Basque words as "prefixes". It seems that, for him, "prefixes" must appear in every word in order to acquire this status. This is a rather strange requirement, as in this case, the well-known Indo-European *s*-"mobile" would not be admitted, as it does *not* (of course) appear before every Indo-European root. Or, to take a

less trivial example, I wonder what would have been Professor Trask's reaction confronted with Meillet's proposal for recognizing two other Indo-European prefixes, namely a *k-* in Latin *caper*, *aper*, *os*, *costa* and for Latin *odium* / Gothic *hattis* (cf. Ernout-Meillet, 1967:38, 95) plus a *d-* for *lacruma* (idem, 336)? Perhaps, respectful for the Master, he would have politely acquiesced, reserving his own negative judgment for private conversations ... Some prefixes were hypothesized for Basque many years before Čirikba and Bengtson, by Schuchardt (1923:5) and H. Berger (1956:16, 1959:24, 30, 36-7); this last paper was read and commented on by Michelena, Berger's "Baskischlehrer und Freund", as mentioned in note 4, p. 18. It must, however, be noticed that Professor Berger no longer holds the theories contained in these papers, which he now considers as "eine linguistische Jugendsünde" — a linguistic indiscretion of youth [personal communication, 11 November 1992].

We may go a little further. Admitting, as a working hypothesis, a connection (either genetic proper, or by *Sprachbund*) between Basque and Sino-Caucasian, it is legitimate to search for Basque *remnants* of Sino-Caucasian prefixes. Then, a cursory look at Benedict-Matisoff *Conspectus* (1972:103-123, 131-133, 154-156) or Shafer's monumental study (1974:20-35) will convince everybody that Sino-Tibetan had a lot of prefixes (but of course not appearing before every word, and not in every language with the same uses). Among them, there is a *b-*, also represented in Chinese, which is hypothesized by Bengtson for a number of Basque words, and an *a-* (only in Tibeto-Burman; where Benedict and Mattisof, p. 123, consider that it was originally "the TB 3rd person pronoun"), parallel to that proposed as a Basque prefix by Schuchardt and Berger, as well as later on by Bengtson. Now, scanning Starostin's "Nostratic and Sino-Caucasian" (1989:106-124) one may easily observe the presence of prefixes in many proposed SC reconstructions, in general as **'V-* , sometimes **HV-* (e.g., in his #18).

Accepting the risk of incurring even more severe denunciatory comments from Professor Trask than those he inflicted on Bengtson et al, I will venture the (far-fetched) hypothesis of a Basque- and Sino-Caucasian plus Burushaski plus Sumerian very remote *genetic* relation with the African superphylum known as "Khoisan". For the last ten years, I have assembled a rather impressive list of putative "related form"' for Khoisan, Sumerian, and Burushaski, part of it entered on Macintosh diskettes. One of the structural elements included is the Hadza prefix *'V / hV* (Ehret, 1986:110) quite similar to the North Caucasic ones and to the *V-* prefixes hypothesized for Basque.

Professor Trask has a very restrictive approach to etymological research. He declares unsuitable for that those roots which he thinks too "universal" to be retained. See his items 62 (Basque *bero* 'hot'), 212-213 (Basque *mama* 'liquid/breast'; 224 Basque *-n* 'locative suffix') or "phonesthetic" (#215 Basque *mara-mara* 'smoothly'; 244 Basque *pinpirin* 'butterfly'; 272 Basque *txiki* 'small') or "imitative" (#269 Basque *tu* 'saliva') or "expressive" (#270 Basque *txaxal* 'calf'). It would be easy to object that "respectable" etymologists, like Meillet (and, later, Watkins) for Indo-European, Burrow and Emeneau for Dravidian, or Rédei for Uralic, did not hesitate to use similar items.

For two of the short list of grammatical "cognates" listed by Bengtson in support of his "Macro-Caucasic" — viz., *hi* 'thou' #145, and *ni* 'I' #237 — Professor Trask remarks that they do not exhibit traces of the kind of alternations present in the Basque forms, and thus fail to provide the nice proof of genetic relations which Indo-Europeanists enjoy finding in similar cases. Certainly, this objection is valid, *if* one accepts the shallow time-depth allowed by Bengtson for his "Proto-Macro-Caucasic" (5000/7000 BP, "comparable to Indo-European" in his words). But, even accepting the Macro-Caucasian hypothesis, one would be extremely skeptical about such a recent date.

Leaving out of account the questionable aspects mentioned above, Professor Trask's detailed screening of those recent proposals for connecting Basque to North Caucasian and some other languages, more or less related to this family, is a fine piece of scholarly research, of a kind which we might hope to see carried out on other proposals for long range etymologies.

What remains positive in the work accomplished by our 'three musketeers'? And the same question for some of their colleagues, past or contemporary, who have tried a more or less similar grouping? Well, some elements are certainly insufficient for validating the membership of Basque in the vast superphylum first proposed by Swadesh as "Basque-Dennean" (or Vasco-Dene — editor), otherwise than at a very deep level. The more restricted macro-family proposed by Bengtson as "Macro-Caucasic" appears falsified, as no specific apomorphisms (innovations exclusively shared by the group, in cladistic terminology) could be detected, using the list in Bengtson, 1990.

Taking Ruhlen, 1990:14, table 4, "Dene-Caucasian cognates", which contains 20 supposed Basque elements, in a list of 35 items, it appears (according to my checkings) that Professor Trask's criticism eliminates six of these (#52, 75, 106, 176, 283, 286 in his list). One more (230 Basque *odol* 'blood') must be cancelled so far as the Burushaski word is concerned (loan from Shina, Turner Nr. 5958) leaving as putative "cognates" only a Na-Dene **del* represented in Eyak and Athapaskan. Nine items (# 83, 134, 166, 171, 191, 229, 253, 269, 302) are not "Dene-Caucasian" apomorphisms, as they appear in other phyla, generally in Khoisan, but also in Old Turkish (Gabain, 1974:345) for #83 and in Carib (Amerind: Greenberg, 1987:316, #104) for # 229; #269 Basque *tu* 'spit, saliva' exhibiting similar forms in Mosan (Amerind, according to Greenberg, but reintegrated in "Dene-Caucasian" if one accepts the change proposed by Shevoroshkin, 1990:9-12), Indo-European, Afrasian, Bantu, and Austronesian (cf. my paper for the first meeting of the Language Origins Society, Krakow, 1975:59-61). One item (#218) corresponds to two different Basque forms: *mihi* < **bini* 'tongue' and *milika-* 'to lick'; the first must be left aside and, in fact, has not been retained by Ruhlen in his Table 4; the second is represented in other phyla than "Dene-Caucasian", for instance Indo-European **melg-*. Thus, we are left with only three items apparently representing Dene-Caucasian exclusive innovations: #120 Basque *gose* 'hungry', 158 Basque *igel* 'frog' (which, if really related to the forms cited by Ruhlen, must have undergone rather drastic and unusual sound changes) and 266 Basque *-t* < **da* 1st singular agreement suffix in verbs (related by Ruhlen with Sino-Caucasian forms for 'we' showing a *t* as initial or medial — a comparison qualified by Trask, not without reason, as "far-fetched"). This very small residue is inadequate for validating the inclusion of Basque in the proposed "Dene-Caucasian" superphylum.

Of course, there are many more proposed correspondences between Basque forms and Sino-Caucasian (without Na-Dene) reconstructions in the 309 different items from Professor Trask's list. As noticed above, their discussion would require much too much space to be inserted here. My provisional conclusion is that there are indeed a certain number of valid correspondences between Basque forms and North-Caucasian and/or Sino-Caucasian reconstructions, but not sufficiently numerous for justifying, at the present stage, the inclusion of Basque into one or the other proposed phylum.

What about "proto-World" (or "Global") etymologies? (I will not discuss here this more or less elusive notion. For me, what our musketeers call, rather inappropriately, a "global etymology" is just "a certain phonemic pattern corresponding to a definable semantic unit, represented in more than one linguistic superphylum"). Professor Trask included in his list eight of them: his #40, 109, 126, 167, 220, 244, 247, and 264. Only one (109) appears really falsified, so far as Basque is concerned. The others would require too much space, again, to be discussed here. Two (167, 220) show a consonant Ablaut similar to that studied by Hodge for Indo-European and Afrasian; if 220 is definitely cognate with Latin *fūnis*, it will exhibit a semantic parallel with Khoisan, where both S1 and N2 show that thoughts were supposed to come from "thinking strings" situated in the throat (see Bleek, 1956:365, 558, 577, 589) — but the forms are there different from the Latin one, which has no accepted etymology (see Walde-Hofmann, I, 567-8; Ernout-Meillet, 262;,Chantaine, 450 under θώμιγξ).

Reviving an old conjecture from Van Ginneken, taken over by Stopa (1960:27-30; 1979:90-91; many details of his scheme now superseded by later research) made more plausible by Cavalli-Sforza et al's biomolecular research (1994), we may venture a paleofiction scenario according to which those now "isolated" fragments of "Mother Tongue(s)" — Basque, Sumerian, Burushaski — were spoken by tribes members of one or several early migrations of "modern Man" taking the "Northern route of expansion" as delineated in Cavalli-Sforza et al, 1994:156 (fig. 2.15.1), where they could have encountered (and possibly married) members of those corresponding to Greenberg's Eurasiatic and Starostin's Nostratic proper (excluding, as in his Cold Spring Harbor paper, 1990:33, Dravidian and Afrasian), as well as others, ancestors of Caucasian (North and South), Yenisseian, and Sino-Tibetan people, maybe also Na-Dene, and even Almosan-Keresiouan. But we must await the results of "fossil DNA" analysis for biomolecular data confirming (or invalidating ...) this bold hypothesis.

References

Benedict, Paul K. and James A. Matisoff. 1972. *Sino-Tibetan: a conspectus*. Cambridge, University Press.

Bengtson, John D. 1990. "An end to splendid isolation: the Macro-Caucasian phylum," *Mother Tongue* (10) April 1990, item 6.

Berger, Hermann. 1956. "Mittelmeerische Kulturpflanzen aus dem Burushaski." *Münchener Studien zur Sprachwissenschaft,* 9: 4-33

——— 1959. *Die Burushaski-Lehnwörter in der Zigeunersprache.*

——— 1974. *Das Yasin-Burushaski (Werchikwar).* Wiesbaden: Harassowitz.

——— Forthcoming. *Die Burushaski-Sprache von Hunza und Nager.*

Bleek, Dorothy F. 1956. *A Bushman dictionary.* New Haven, American Oriental Society.

Burrow, T. and M. B. Emeneau. 1984. *A Dravidian Etymological Dictionary.* 2nd ed. Oxford, Clarendon Press.

Cavalli-Sforza, L. Luca, et al. 1994. *The History and Geography of Human Genes.* Princeton University Press.

Chantraine, Pierre. 1968-77. *Dictionnaire étymologique de la langue grecque.* Paris: Klincksieck.

Ehret, Christopher. 1986. "Proposals on Khoisan reconstruction". *Sprache und Geschichte in Africa,* 7,2: 105-130.

Emeneau: see Burrow and Emeneau.

Ernout, A. and A. Meillet. 1967. *Dictionnaire étymologique de la langue latine.* 4th ed., Paris, Klincksieck.

Gabain, A. von. 1974. *Altürkische Grammatik.* Wiesbaden, Harrassowitz.

Greenberg, Joseph H. 1987. *Language in the Americas.* Stanford, Stanford University Press.

Grolier, Eric de. 1985. "Notes on the methodology for possible linguistic reconstructions beyond the limit of 10,000-15,000 years BP." Paper for the first annual meeting of the Langauge Origins Society, Mogilany (Krakow).

Hodge, Carleton T. 1986. "Indo-European consonant ablaut." *Diachronica* 3:143-63.

Lorimer, D. L. R. 1935-1938. *The Burushaski Language.* 3 vols. Oslo, Institute for the Comparative Study of Civilizations.

——— 1962. *Werchikwar-English Vocabulary.* Oslo, Institute for the Comparative Study of Civilizations.

Meillet, Antoine: see Ernout and Meillet.

Nikolaev, S. and S. Starostin. 1990. "North-Caucasian roots; ed. and arranged by A. Eulenberg," in Vitaly Shevoroshkin, ed., *Dene-Sino-Caucasian languages,* pp. 174-264. Bochum: Brockmeyer.

Rédei, Károly. 1986-. *Uralisches etymologisches Wörterbuch.* Wiesbaden: Harrassowitz.

Ruhlen, Merritt. 1990. "An overview of genetic classification," in J. A .Hawkins and M. Gell-Mann, eds., *The Evolution of Human Languages.* Addison-Wesley.

Schuchardt, Hugo. 1923. *Primitiae linguae vascorum.* Halle. Reprinted. Tübingen, Niemeyer, 1968.

Shafer, Robert. 1974. *Introduction to Sino-Tibetan.* Wiesbaden, Harrassowitz

Shevoroshkin, Vitaly. 1990. "Introduction," in Vitaly Shevoroshkin, ed., *Proto-languages and Proto-Cultures,* pp. 8-12. Bochum, Brockmeyer.

Starostin, Sergei A. 1989. "Nostratic and Sino-Caucasian," in *Materialy k diskussijam no Mezhdunarodnoj konferencii (Moskva, 29 Maja-2 ijunja 1989g.),* pp.106-124.

——— 1990. "A statistical evalution of the time-depth and subgrouping of the Nostratic macrofamily," in *Abstracts of Papers Presented at the 1990 Meeting on Evolution: from Molecules to Culture.* p. 33. Arranged by Richard Dawkins and Jared Diamond. Cold Spring Harbor Laboratory.

Stopa, Roman. 1960. *The Evolution of Click Sounds in some African Languages.* Krakow: Nakladem Uniwersytetu Jagiellonskiego.

Swadesh, Morris. 1971. *The Origin and Development of Language.* Chicago, Aldine.

Turner, Ralph Lilley. 1966. *A Comparative Dictionary of the Indo-Aryan Languages.* London.

Walde, A. 1965. *Lateinisches etymologisches Wörterbuch.* 4. Auflage (hrsg. J. B. Hofmann). Heidelberg: Winter.

Watkins, Calvert. *The American Heritage Dictionary of Indo-European Roots.* Boston, Houghton Mifflin Company.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

COMMENTS ON LARRY TRASK'S PAPER

José Ignacio Hualde
Department of Spanish, Italian, and Portuguese
University of Illinois at Urbana-Champaign
4080 Foreign Languages Building
707 South Mathews Avenue, Urbana, IL 61801
Telephone (217) 333-3390; fax (217) 244-0190; e-mail sip@uiuc.edu .

[Editor's Note: Professor Hualde is a native speaker of Basque and also a linguist, recommended as a scholar broadly supportive of Trask's position and an internationally respected scholar on matters of Basque linguistics and taxonomy. Because of a flaw in our communications, his comments are addressed to Trask and this paper is a copy of his e-mail message to Trask. For appearances sake, we have modified his text slightly from its e-mail format. All references 'p' are to page numbers in Trask's text. Numbers in brackets refer to the etymologies numbered by Trask. Page numbers refer to Trask's original manuscript.]

p.1: Basque-Iberian. "... and hence that Iberian could not be an ancestral form of Basque." But consider Gomez-Moreno's position, who maintains that Basque is an Iberian dialect. He presents two arguments. On the one hand we have the toponyms *Il(l)i-* which are found all over the Peninsula and which is sensible to relate to Basque *iri* 'town'. There is also Selaya near Santander (cf. Basque *zelai* 'meadow'), Iturissa < Tossa de Mar, near Girona (cf. Basque *iturri-tza* 'place of springs' vel simile), etc. For Gomez-Moreno, a second, very strong, argument is provided by the famous Bronze tablet of Ascoli. This table contains names of equites from places such as Ilerda, which we know were Iberian-speaking, and also equites from the territory that we know belonged to the Vascones. As Gomez-Moreno points out, there is no distinction between Basque and Iberian names in this tablet. All names have the same structure, and many Iberian onomastic components appear in the name of equites from the territory of the Vascones. Gomez-Moreno also points out that it is sensible to assume that all the equites in this military unit spoke the same language. That is, the Vascones were ethnically and linguistically Iberian. The fact remains that modern Basque is not of much use in deciphering the ancient Iberian inscriptions. But to this, Gomez-Moreno replies that the Celtiberian inscriptions are equally opaque to him, even though they are obviously written in an Indo-European language. For him, this is not a definite argument. As he remarks, a modern Romance dialect (e.g., French or Milanese), probably would not help us much in the decipherment of an inscription in archaic Latin. To this we must add the fact that there would be some difference between the Iberian of the Mediterranean coast (where most of the inscriptions have been found) and the Iberian dialect that the Vascones spoke. These arguments seem sound to me. In my opinion, the truly damaging fact for the Basque-Iberian hypothesis is not that we cannot read the Iberian inscriptions by means of Basque, but the stark contrast provided by the transparently Basque Aquitanian inscriptions.

p.5, last line: *antzara* > *antzar*
The list on pages 5-6 includes both unquestionable loanwords and words whose etymology is less clear (*izokin*, *haizkora*) even though they *could* be borrowings.

p.6, paragraph 2: Add Navarrese Romance. Delete Bearnais.

p.7, last paragraph: "no native word begins with any voiceless plosive." This is too strong. Off the top of my head I can think of a few examples: *ke* 'smoke', *koipe* 'grease', *kirats* 'stench'. In a couple of hours, I (or you) could come up with fifteen or so such words which are not obvious borrowings.

p.8: "Latin initial voiceless plosives are consistently rendered by voiced plosives." Again this is too strong. What one finds is a certain amount of fluctuation of the type *bake* ~ *pake* 'peace'. Martinet and Michelena assume that the forms with a voiced stop are older, but this is only a hypothesis. The fluctuation extends to modern borrowings. Michelena gives examples for his native variety of Renteria somewhere. I believe the examples he

gives are *gortiña* 'curtain' < Spanish *cortina*, *gorbat*a 'necktie' < Spanish *corbata*, but *kabarra* 'type of boat' < Spanish *gabarra*. (I do not have the reference right now, but I could find it if you want me to.) He indicates that, within a given variety, one consistently finds a voiceless or a voiced stop in a given item, but dialects differ in the particular solution adopted for a given word.

Your statement that "Basque words beginning with voiceless plosives [...] are thus ALWAYS loan words or recent formulations" is a hypothesis, not a fact.

p. 10: Under **ardano* you may wish to add Latin *arno*.

p.11: Concerning *onhore* > *ohore*. This is not the only possible derivation. Notice that what we have in Z is a nasalized /*h*/, i.e. [*h*~] which causes the nasalization of neighboring vowels in the same way the other nasal consonants do. A possible evolution is /*n*/ > [*h*~] // V__V (i.e., deletion of oral cavity features).[1]

p.11 again: True, there is a certain amount of variation regarding the distribution of /*h*/ in different varieties. But this variation is not so overwhelming. Otherwise it would have been impossible to determine the use of <h> in standard Basque orthography. Michelena's hypothesis regarding accentuation also assumes a high degree of consistency in the distribution of /*h*/.

p.11 again: Your first constraints on the distribution of aspiration has to be reformulated. The aspiration cannot follow a word-initial liquid or nasal: **lhabur*, **nhoiz*.

I would be a lot more impressed, though, if somebody could find a couple of Basque grammatical morphemes in Iberian. I know it has *-ko*, but Indo-European has *-ko*.

Well, it (Iberian — editor) has *-taR* ~ *-daR*, seemingly with the same distribution (and meaning) as Basque *-(t)arr*, and Michelena has an article on "Iberico *-en**" (a genitive).

[2] *abere*: Bengtson, in his reply, seems to think that this could not possibly be a borrowing from Latin *habere* (an infinitive). He notes the existence of Spanish *haberes* 'possessions', but he is not impressed since the Spanish word does not mean 'animal'. I do not see the problem. A parallel case is given by the HN (quasi)-synonym *azienda* 'livestock'. Notice that Spanish *hacienda* (from *hacer* 'to make') means 'finances' and 'estate'. Depending on the area a man's most important possession can be his land or his livestock.

[3] [Basque *abets* 'voice' — ed.]: Sure, *abets* is out. But notice that *ahots* alternates with *abots* 'voice', since the Abkh parallel that is given has a schwa, I suppose the comparison is still valid (you and I know that *abots* derives from *aho* 'mouth' + *hots* 'sound' and that the /*b*/ is also found in *ahoa* ~ *aba* 'the mouth', but never mind).

[4] [Basque *adar* 'horn, branch' — ed.]: What you say about Irish does not destroy the comparison. It could very well be a proto-Basque-Caucasian word borrowed into Irish.

[10] [Basque *ahizpa* 'sister (of a woman) — ed.]: Bengtson incorrectly translates *ahizpa* as 'sister (woman speaking)'. The only correct translation is the one you give above — 'sister (of a woman)'.

[11] [Basque *aho* 'mouth' — ed.]: You say: "As always, Basque *h* is not etymological." Too strong, if you mean that it is always a 'prosodic feature'[2]. As you know, there are some cases where it continues some earlier segment: the demonstratives and related adverbials *hau*, *hori*, *hura*, *hemen* ... (cf. Renteria *kau*, *kori*, *kura*, *keben* and Sal, Aezk, etc. *gau*, *gori* ...), possibly *harri* 'rock' and a few others.

[1] Editor's note: for non-linguists, his symbols here mean that an 'n' changes to a nasalized 'h' when it occurs between two vowels. Z refers to an eastern dialect. Trask actually refers to 'Honore' > 'ohore'.

[2] Editor's note: for non-linguists, the issue here is whether an [h] occurs only for reasons of (what we used to call) euphony or sounding good or perhaps looking right (in writing), rather than being a true phoneme of Basque origin.

[27] [Basque *arrain* 'fish' — ed.]: It would be only fair that the reconstructed form is **arrani*, which brings it closer to the Coptic form. [Editor's note: Faulkner gives Middle Egyptian */rm/* simply as 'fish'.]

[38] [Basque *azal* 'skin, bark' —ed.]: You say: "moreover an initial **k* points indisputably to a loan word." Not indisputably.

[47] [Basque *barrabil* 'testicle' — ed.]: The words *barret* and *barrasaket* are unknown to me, but they look very Gascon-like. The final *-t* is rather un-Basque.

[49] [Basque *baso* 'woods' — ed.]: Your comments about the "predominantly deciduous Basque forests" are true for most of the present-day Basque-speaking area, but not for other formerly Basque-speaking regions. In the Pyrenees conifers are more common than other trees. Renteria preserved the word *ler* 'pine tree', whereas the borrowing *pinu* is found elsewhere. The vegetation found on the Aquitanian plains is also rather different from that of the modern Basque Country.

[52] [Basque *beginini(ko)* 'pupil (of the eye)' — ed.]: No way Spanish *niño* is a borrowing. Similar forms are found in other Romance languages: Portuguese *menino*, Bal., Catalan *nin* (other Catalan dialects have *nen*). An expressive origin is very likely, as you point out. Cf. Italian *ninna nanna* 'lullaby', *ninnare* 'to rock a cradle'.

[55] [Basque *beko* ~ *moko* 'forehead; beak, extremity' — ed.]: You may want to add that *ikusi-makusi* follows a common pattern of 'expressive' reduplication: *haundi-maundi*, *duda-muda*, *inguru-minguru*, etc.

[60] [Basque *berezi* 'separate' — ed.]: Since the root is *ber-* in any case, the comparison is still valid.

[74] [Basque *bizi* 'alive' — ed.]: In spite of what you say, *bizi* does behave like a (an irregular) verb in probably most varieties: *biziko da* and not *bizi izango da*. The irregularity which indicates its adjectival origin is that *bizi da* is a present and not a present perfect, but *heldu* behaves in the same way in many areas: *heldu da* 's/he is coming', even though it has the same ending as *galdu*, etc.

[98] [Basque *esne* 'milk' — ed.]: Michelena reconstructs **esende*, which happens to offer a much better match. It is only fair also to point out the cases where the reconstructed Proto-Basque form helps Bengtson.

[103] [Basque *ez* 'not, no' — ed.]: The form *ze* is still used in the subjunctive in some Bizkaian varieties. Gaminde gives some examples from Berango such as *koxi señe yausi sedeitten, imiñi erloju loiñ sedeigun.*

[113] [Basque *giltza* 'key' — ed.]: About *giltza*. For the connection between the meanings of 'key' and 'nail', cf. Latin *clāvus* 'nail' and *clāvis* 'key' (which give Catalan *el clau* 'the nail' and *la clau* 'the key', for instance.)

[114] [Basque *giltzurrin* 'kidney' — ed.]: As far as the second element *urdin* is concerned, you say it was "formerly 'green'"? Is this right? Do you mean 'gray'? It still means 'gray' in some areas.

[117] [Basque *gogor* 'hard' — ed.]: For the semantic connection between 'hard' and 'deaf' see 'hard of hearing' (Spanish *duro de oido*). There is a form *entzungor* (consisting of *entzun* 'hear' + *gor*).

[119] [Basque **gor; gorri* 'red' — ed.]: You may add *zuringo*, 'white of an egg'.[3]

[3] Editor's note. For non-linguists, the complexity of these arguments may be simplified a bit by observing what we have just learned, to wit, *gorri-ngo* and *zuri-ngo* mean respectively 'red of egg' (= yolk) and 'white of egg'.

[123] [Basque *(g)une* 'stretch, interval, moment, occasion', occurring as *-gune* and *-une* in various contexts. — ed.]: You say: "there are many parallels in Basque for the loss of initial *g-*, but hardly any for the insertion of an initial *g-* ." An example is BN *garratoi* 'rat', compared with common *arratoi* (from *raton(e)*).

[125] [Basque *gurpil* 'wheel', *inguru* 'vicinity, around, near' — ed.]: Bengtson argues that, among others, *in gyrum* is too learned a word for the "crusty Bizkaian shepherds and farmers" to borrow. Bengtson continues: "Similarly Trask derives Basque *inguru* 'vicinity, around, near' from Latin *in gyrum* or *in giru* 'in a circle', again ultimately of Greek origin. But how did it get from here to there?" (Page 8, presumably of Bengtson's work) I don't see the problem. Notice that *giro, andare in giro* is perfectly colloquial Italian.

[135] [Basque *hau* 'this', *han* 'over there', *hara* 'thither' — ed.]: This is a good point to discuss the *k-* or *g-* initial forms of the demonstratives and related adverbs in several dialects.

[139] [Basque *hazi* 'grow, grow up, raise, bring up, cultivate' — ed.]: In support of the parallel with Abkhaz *a-z-ha-ra*, notice that Michelena derives *azkar* 'strong' from *(h)azi* 'grow(n)'.

[156] [Basque *ibar* 'water-meadow, valley' — ed.]: Your geographical observation that "In the Basque Country, with its narrow V-shaped valleys, such terrain is found nowhere but in valleys" is not valid for Aquitania or the southern Alavese and Navarrese plains.

[194] [Aquitanian **kala* — ed.]: About your statement that "Aquitanian, of course, doesn't even have any stems beginning with *k*." !!?? But see Cisson.

About Calagurris, there were several towns with this name, as Gorrochategui points out in his paper.

[250] [Basque *-ra* (allative case-suffix) — ed.]: But Renteria has *-ara*, for example, *etxe-ara* 'to the house'. Cf. also Z. *-ala(t)*.

[300] [Basque *zazpi* 'seven' — ed.]: This is one of Schuchardt's brilliant matches. H. Pedersen (*The Discovery of Language*, p.127) remarks that "It is a little awkward that the resemblance is greater to Coptic than it is to Old Egyptian, comparison with which shows that the two last consonants in Coptic *sašf* originally occurred in reverse order; so that if the relation is to be maintained, we must also accept an inversion of the consonants in the Basque *-zp-* ."[4]

[302] [Basque *zer* 'what?' —ed.]: Bengtson deserves some credit here. Cf. the alternation between western *zela(n)* and central and eastern *nola* both meaning 'how'.

[304] [Basque *zigar* 'mite' — ed.] I wonder if Spanish *cigarra* 'cicada' has something to do with this word.

[308] [Basque *zilar* 'silver' — ed.]: To Agud and Tovar, pace! But *silab(u)r* is also found in Celtiberian (Bronze inscription of Botorrita). But since the expected form in a Celtic language would be **argantom* vel sim., as Michelena points out, the direction of the borrowing remains unclear.[5] What is clear is that the Celtiberian attestation shows that the Basque word probably predates the Germanic invasions by a long shot.

---

[4] Editor's note: In Pyramid texts and later Old Egyptian had *sfx-w/-t* for 'seven'. If we add the schwa, but retain the Basque pattern, then it would have been something like **səfxi* or **səfHi* .

[5] Editor's note. The expression 'vel sim.' is from vel simile and roughly means 'or the like, something similar, a like value'.

One final comment: [He is referring to Trask's discussion of case-endings on page 99 of his original text — editor]. You say: "Basque *-ta* is not a case-ending at all, let alone a 'locative' (as Bengtson would have it), and it is moreover probably of Latin origin." I don't think Bengtson is so wrong here. First, it is true that *-ta* is only found as a suffix in non-singular locative forms (inessive, allative, ablative, and so-called 'genitive locative' — not strictly a locative form, but used for the most part with place NPs). Second, there are serious problems for a Latin etymology. I could accept this if we had only plural *-eta* (for example, *mendietan* 'in the mountains', *mendietara* 'to the mountains', etc.) but there is also indefinite *-ta* (for example, *menditan* 'in mountain(s)', *menditara* 'to mountain(s)', etc.). What is the source of the indefinite locative suffix?

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

COMMENT ON R. L. TRASK'S "BASQUE AND DENE-CAUCASIAN: A CRITIQUE FROM THE BASQUE SIDE"

William H. Jacobsen, Jr.
University of Nevada

1. Introduction

Given my longstanding interest both in Basque historical linguistics and in the characterization and evaluation of evidence for distant language relationships, I was glad to be asked to contribute comments on this impressive paper. I am doing this under tight time constraints, however, as I was invited much later than the other commentators; I might have welcomed an opportunity for more ruminative consultation of the literature, especially that concerning previous comparisons to Caucasian languages, potential Latin-Romance etymologies, and patterns of semantic shifts. I have not been provided with copies of the original publications from which Trask has extracted these comparisons, and so may not be aware of additional explanations, bibliography, etc. that they may contain. (I also have not had available the Spanish etymological dictionary of Corominas and Pascual [1980], so I have cited the much earlier Corominas 1954, checked against the more recent but greatly abridged Corominas 1973.) John Bengston has kindly provided me with a copy of his comment, so I have taken into account his comments on individual items and on some phonological issues, but it is not my assignment to deal with the new material that he introduces there.

Workers in this field should be greatly indebted to Trask for all the effort he has expended in trying to make sure that Basque data is accurately handled in long-range comparisons. It is rather rare for experts in a little-known language to contribute data to already on-going comparative efforts. In the field of Hokan studies, I am reminded of Waterhouse (1976), who offered corrections, and especially additions, to several comparative studies based on her knowledge of Chontal of Oaxaca (or Tequistlatec), and of Jacobsen (1979:566-70), who contributed Washo corrections and data to the ambitious study of Gursky (1974). And of course a veritable cottage industry has emerged for presenting corrections to, or at least complaining about, the data in Greenberg 1987. So where are you, o Burushaski expert?

A few comments relating to Trask's §1. There seem to have been as many as five different linguistic groups in the Iberian Peninsula based on the earliest attestations. I would especially recommend Hoz (1981) as a useful summary of the situation. To speculate a little: Gensler (1993) offers a sophisticated attempt at demonstrating a "Mediterranean Hamito-Semitic" substratum to Insular Celtic: could one of the southern or western peninsular groups represent this population in an intermediate location?

For work on Basque-Caucasian comparisons in recent decades, one should also mention Jan Braun (1981), who compares Basque to Kartvelian (South Caucasian), showing good control of the Basque data and attempting to establish sound correspondences. Urreiztieta-Rivera (1980) is a useful survey of Basque-Caucasian comparative work, perhaps more descriptive than critical. I would also like to call attention to Kuipers (1960:109-10), who starts out from a good analytical knowledge of Northwest Caucasian to criticize six comparative sets of Bouda (1948, 1949). Some additional work by these authors that seems worth mentioning is Bouda 1952b, Bouda and Baumgartl 1955, Lafon l947, and Uhlenbeck 1942. Uhlenbeck 1949 is a more accessible translation of Uhlenbeck 1927.

In his §4, Trask offers a summary of the characteristics and nature of the evidence in these sets of comparisons, with which I am in general agreement. However, specificity is needed in these matters, and it seems that one contribution I might make is in bookkeeping, in independently judging and keeping track of exactly which and how many comparisons have defects or limitations of various kinds. We start out from the following situation. Of the 317 numbered sets of comparisons, 10 are merely cross-references to separate words occurring in other sets; another pair should be merged: [191.] **-k** 'thou' with [145.] **hi** 'id.', as not offering additional forms or comparisons. This leaves us with 306 sets of comparisons as brought together by Trask, many of which combine, of course, the contributions of two or more different authors (most commonly Čirikba with Bengston).

Conversely, several of the sets bring into comparison what I take to be two or more different Basque words or affixes, as follows: 2 each in [7. 34. 39. 83. 125. 170. 193. 194. 217. 241. 259. 260. 278. 284. 296. 301. 310. 317.], 3 each in [55. 119. 171. 264. 311.], 4 in [239.], and 6 in [135.]. This increases the number of Basque (or Spanish!) words, stems, or affixes compared to 342, at first approximation. (A few words recur in more than one set, but I count them only once, in the first set listed: *bizkar* 'back' [75. 288.], *t(h)u* spit' [269. 171.], *ukaiṅ* 'wave' [280. 159.], *-ik* (partitive suffix) [160. 190.], and *-n* (locative suffix) [224. 123.]). For more technical reasons, as discussed individually, yet a few additional cases of the recurrence of the same Basque morpheme in different comparisons can be noted: *ni, n-* 'I, me' [227.] and *-t ~ -da-* 'id.' [266.], *azkonar* 'badger' [40.] and *hartz* 'bear' [132.], *urtxakur* 'otter' [286.] and *zakur, txakur* 'dog' [296.], *hau(r)* 'this' [136.] and *hau* 'id.' [135.], and *-a* (article) [1.] and forms of *har-* 'that' [135.].)

For the sake of systematic discussion, however, we need not regard those sets of comparisons with three or more members as single packages, to be accepted or rejected as wholes. Rather, each separate pairing of a Basque form with one from another language or family constitutes a separate hypothesis that can be separately discussed and evaluated. For example (to pick on some insects), when focusing just on meanings, we can note that the comparison of Basque *sist* 'moth' [261.] with a Sumerian word for 'moth' is a better match than with a Proto-North-Caucasian word for 'bug, tick' or a Burushaski word for 'a kind of beetle', or again, that Basque *euli* 'fly' (n.) [101.] matches the (Northeast Caucasian) Archi word for 'fly' better than the Burushaski word for 'butterfly, moth'.

Thus I will speak of *(comparative) sets* on the one hand, meaning usually the forms gathered together by Trask in a numbered entry, but sometimes just those of a single earlier author, and of *(comparative) pairs* on the other, which should be the meaningful units for discussion and evaluation. Recognizing these pairs brings about, in this case, an approximate doubling, from 342 Basque words to 650 comparative pairs. In a few sets, we have a marked increase due to multiplication; for example, [7.] compares 2 Basque words for 'dry' to words in 3 other languages, making 6 pairs; while [171.] compares 4 Basque words with different meanings, 'drop', 'spit', etc., to words in 3 other languages, making 12 pairs. In determining the number of pairs present, I have tried to tilt towards assuming that North Caucasian forms represent a single proto-form.

Some of the sets contain comparisons to Proto-Dene-Caucasian or Proto-World reconstructions, which rather confuse the picture, as these are not comparisons to specific sister forms, but rather to shorthand formulas intended themselves to justify more wide-ranging comparisons. However, such sets are fairly few in number: there are just 4 with Proto-Dene-Caucasian reconstructions [83 (with 3 reconstructions). 97. 109. 295.], and 9 with Proto-World [40. 109. 126. 167. 220. 244. 247. 264. 297.].

In advocating comparative pairs as the unit of reckoning, I should not be misunderstood as rejecting the principle that a higher average number of members in sets, other things being equal, potentially makes a relationship among three or more words, or languages, more likely, a point taken up near the end of §6.

I will proceed by considering several relevant questions of Proto-Basque consonantism in §2, discussing morphological segmentation and a suggested derivational prefix in §3, and commenting on many of the comparative sets in §4, and then returning to the bookkeeping, for affixes in §5 and words and stems in §6.

2. Phonology

Three generalizations made by Trask in his §2 raise questions that require fairly extensive discussion: that there was no **m* in early Basque [3], that no native Basque word begins with a voiceless stop *p, t,* or *k* [1], and that *h* does not reflect a consonantal segment of early Basque [7]. These are interrelated questions, most closely the first two, i.e., nasals and stop series. Negative generalizations like the first two are inherently risky, of course, since it is perfectly possible that a phoneme might be of rare occurrence, either in general or in a specific environment, and yet not be completely absent. A good example is found in Trask's additional, and appropriate, generalization that also *d* does not begin native words. In connected texts, there occur, in fact, numerous native words beginning with this consonant, but they are all present tense finite verbs beginning with the third person absolutive prefix *d-*. (I give a tentative explanation for this exception in my comment, §4 [80.].)

In discussing **m*, I will consider first the initial, and subsequently the medial, environment. In my comment on first person singular *-t* [266.], I point to evidence explaining the absence of *d-* as the result of a general

sound change of **d-* to *n-*. In parallel fashion, I would like to suggest that what is actually missing is early **b-*, this having merged with **m-*. Contemporary *b-* comes from **p-*, and *p-* is a secondary introduction.

Exactly these adjustments, *m-* and *b- > m-*, *p- > b-*, are seen as the predominant pattern in the early Latin borrowings. This is probably not because they came in early enough to get caught up in the **b- > m-* change, but because then there were just two bilabial phonemes in initial position, **p-* and **m-*, so that by virtue of being voiced Basque *m-* would have been equated to two Latin consonants. Michelena (1977:275), after Voegelin and Voegelin (1959:13), suggestively points to a phoneme in (Siouan) Crow and Hidatsa that contrasts with /p/ and has three allophones, [w], [m], and [b]. Especially since it may result from a merger of two phonemes, I see no clear evidence that **m-* was absent from early Basque. And it is instructive to note that, in order to explain away some Basque nouns starting with *m-*, Michelena (1977:271-73) recognizes a meaningless nominal prefix *m(a)-* — exactly the sort of thing Trask (§4) objects to when it comes to *b(e)-*! Thus I am not convinced that this assumption suffices to exclude from being native Basque several words in our sets: *moko* 'beak, front, face' [55.], *moto* 'headscarf, etc.' [217.], *muin* 'pith, marrow, inner part' [220.], and *mutur* 'snout, muzzle' [223.].

The assumption that early Basque had no **m* in medial position seems to be based on the plausible equating of Aquitanian *SEMBE* with Basque *seme* 'son', along with the existence of unquestioned processes for deriving *-m-*, especially by nasal infection of *-b-* or *-w-*. This also seems to me an overgeneralization, and I see no justification for claiming that all cases of *-m-* in native words originate from some other consonant or cluster. Many occurrences are indeed secondary, in nursery words such as *ama* 'mother' [18.], phonesthetic words such as *mara-mara* 'smoothly' [215.], borrowings from Latin such as *eme* 'female' [88. 89.], and dialectal variants such as *imiñi* 'put' [167.] containing a following nasal. Several words in these sets are unquestionably native, and may have had the same sound in early Basque: *amets* 'dream' [19.], *hamar* 'ten' [128.], *-(k)ume* 'child, offspring' [88.], perhaps *ametz* 'gall oak' [20.], and even *samin* 'bitter' [255.]. It might be remarked that in the standard spelling, occurrences of *-np-* and *-nb-*, as in *pinpirin* 'butterfly' [244.] and *sunbin* 'nostril' [264.], conceal occurrences of clusters *-mp-* and *-mb-*, both phonetically and phonemically, exactly as in Spanish.

The bilabial phoneme that really *is* missing from native Basque words is *p*, that would have come from nonexistent early **p^h*. Trask notes this for initial position, as part of the pattern involving all three voiceless stops. Gavel (1921:316-25) gives a thorough discussion of its secondary occurrences initially. In our sets, there are eight words beginning with *p*, of which the only two that seem likely to be native Basque are *p(h)artz ~ patz* 'louse, nit' [48.] and *pintza* 'membrane' [245.]; they have variants with *b-* and *m-* respectively, which are apt to be the original sounds.

It is strange that Trask did not recognize the fact that also medially *p* is not present in native words. In these sets, there are just six occurrences. Two of them are certainly from Latin-Romance: *apo* 'toad' [22.] and *pipil* 'bud' [246.]. In another two, the *-p-* arises regularly from **-tb-* at a morpheme boundary: *bep(h)uru* 'eyebrow' [78.] and *gurpil* 'wheel' [125.]. In another, the *-p-* is placed internally by phonesthetic reduplication: *pinpirin* 'butterfly' [244.]. And the last, the native word *ipiñi* 'put' [167.], contains the verb prefix **e-* and shows variants with *-b-* and *-m-*, so that **b* was doubtless the original consonant.

The second generalization, that no native Basque word begins with *p*, *t*, or *k*, raises two separate issues: 1) whether the occurrence of these sounds flags a non-native word, and 2) the number of stop series present in initial position in early Basque. The answer to the former, as an empirical matter, is clearly yes. Initial *p* has just been discussed. See also Gavel (1921:365-82, 428-42) for a thorough treatment of words starting with *k* and *t*. In our sets, there are just three words starting with *t*: *t(h)u ~ ttu* 'spit, saliva' [269. 171.], which is onomatopoetic, and *tiña* 'ringworm' [170.] and *toska* 'kaolin', Z. *toxka* 'clod' [268.], which are from Latin. There are 11 sets here with words starting with *k*, but the only ones that seem likely to be native, such as *karats* 'bitter' [197.] and *korotz* 'dung' [199.], have variants with *g-*, which is probably the original sound. Variations such as in *ume* 'child, offspring' beside *-kume* as in *emakume* 'woman' [88.] may be attesting to the loss of former **k-* (cf. Michelena 1977:244-6).

Turning to the question of the number of stop series that occurred word-initially in early Basque, there certainly are languages, such as Finnish and Southern Paiute, that are of the type Trask points to, with just one series initially, but two or three medially. Martinet (1950, 1955) thought of two series, with a correlation of aspiration rather than voicing, the aspirated ones having subsequently been lost. My own theory, which I present here without justification, is that early Basque, like Ancient Greek and Thai, had three series, such as **d*, **t*, **t^h* (but no **p^h*), and these survived as such non-initially in Northern dialects in the face of the onslaught of Indo-European languages

with only two series. Initially, a drag chain series of changes took place, **d-* > *n-*, **t-* > *d-*, **tʰ-* > *t-*, so that contemporary initial aspirated stops represent a reintroduction into this position.

Last but not least, we turn to the question of *h* and its possible antecedent consonants. I start out by noting that Michelena (1950c:258-59; 1977:205-07, 219-21), in contrast to Trask, clearly assumed that a phoneme **h* was found in Proto-Basque. This reminds us that it is really inappropriate to speak, as Bengston does, of a "Michelena-Trask Pre-Basque Phonology". Rather there is a separate Michelena phonology, Trask phonology, Jacobsen phonology, etc., sharing, of course, a large core of agreed-upon findings. Moreover, Trask's account unintendedly makes it seem as though Michelena's massive achievement came *ex nihilo*, whereas he freely refers to and builds on work on Basque etymology and reconstruction of other scholars, such as Schuchardt, Uhlenbeck, Gamillscheg, Lafon, Gavel, and Martinet.

Under the label "aspiration" Trask (§2[7]) groups together what are, historically speaking, several different entities so as to claim that *h* (almost) never continues an earlier consonant. To deal with this systematically, we must consider individually different classes of environments. First, we should take seriously the thought behind the orthography for Unified Basque (Euskara Batua), and set aside the postconsonantal occurrences, where it probably does not usually (but may sometimes) represent an older consonantal segment. This leaves us with the initial and intervocalic environments: I think that in the former *h* directly continues, and in the latter indirectly testifies to, a consonant of early Basque.

We take up first the initial environment. There was apparently a time when, as is true for very many languages, such as English, Latin, and Greek, *h* occurred only word-initially (or stem-initially, allowing for prefixes). This consonant was simply present in Proto-Basque, and its origin cannot be confirmed within Basque itself, although it may in part come from earlier aspirated (or at least voiceless) stops, *kʰ-*, *tʰ-*, *pʰ-* (cf. Martinet 1950:228-31, 1955:379-82), and there are some linkages to be seen with *g-* (perhaps < **k-*; cf. my comment on *hi* [145.]).

In these data, there are 83 comparative pairs in which the Basque word starts with *h*. In 30 of these, surprisingly enough, there is no corresponding consonant in the other word compared. For 11 of these, there is no match for a whole initial syllable, whether or not explicitly segmented in the Basque form: shapes *ha-*, *han-*, *he-*, and *ho-*. For the other 19, the matching includes an initial vowel, but no preceding consonant. There are 7 pairs involving Sumerian, where the initial is *Ø-*, *h-*, and *p-* ~ *b-*. With Burushaski, there are 17 pairs, with initials *Ø-*, *h-*, *k-*, *k(h)-*, *g-*, *x-*, *γ-*, *q-*, and perhaps *-y-*. In the comparisons to Caucasian, there are 30 pairs, with at least 15 different consonants (I tried to avoid counting separately reconstructed consonants that subsume other consonants, but do not completely control this). If thought to all stand behind Basque *h-*, this would indeed be multiple origin with a vengeance!

Later this *h-* was joined by an *-h-* that appeared when there was a hiatus between the first and second syllables. Hiatus in this context means contiguous vowels other then the falling diphthongs *ai*, *ei*, *oi*, *au*, *eu*. It is clear that there were two distinct periods of consonant loss here. In the more recent one, it was an *-n-* whose loss has left the hiatus. The *-h-* here does not directly continue the lost consonant, but has become introduced into the hiatus, which would not arise in native words otherwise than by the loss of a consonant, although it might be presented ready-made in Latin words such as *leōne* 'lion'. Thus from Latin *honōre* 'honor' there came (attested) *oore*, from which *ohore*, but there never was the **onhore* suggested by Trask.

The evidence for the former occurrence of **-n-* in a given word can be of several kinds: nasalization of vowels in Zuberoan, but more often present in its now-extinct sub-Pyrenean counterpart Roncalese; independent or combining forms ending in *n*, contingent on loss of the following vowel; and less often, the testimony of earlier records. Since the comparisons made to other languages have mostly not allowed for the occurrence of a nasal consonant, I have surveyed the evidence, and indicate its nature in my comments on those individual items for which Trask had not already done so. One of the sets clearly involves a borrowing from Latin, *xahu* 'clean' [289.]; another borrowing from Latin, *biga* 'two-year-old heifer' [68.], does not show *-h-* because the **-n-* began the third syllable; yet another word, *haitz* 'rock' [127.] may also have had an intervocalic nasal which did not yield *-h-* and therefore did not enter into the comparison (the syllable *ha-* is not matched). Two of the sets for which Trask assumes **-n-* [159. 313.] seem to me to have had a different consonant, belonging to the older period. This leaves us with 10 comparative sets (including *xahu*), involving 15 pairs [9. 10. 12. 85. 218. 232. 270. 289. 301. 305.]. For

one of these, *ahuñe* 'kid' [12.], the evidence for **-n-* comes exclusively from an older form of a toponym; for another, *ohoin* 'thief' [232.], I have suggested an etymology which would confirm the *-n-* but ruin the comparison.

In the comparisons made, there is a straightforward matching to a consonant appropriate to **n* only in one set: *mihi* 'tongue' [218.] with an *-l-* in Proto-North-Caucasian and Burushaski forms, given that *-n-* and *-l-* are seen in potentially related *mintzo* 'voice, speech, conversation' and *milikatu* 'lick'. Two other matchings seem fortuitous: *xahu* 'clean' [289.] with *-r-* ~ *-l-* in a Proto-Yeniseian form, based on an erroneous Basque form with **-nh-*, and **ziho** 'tallow' [305.] to a Proto-North-Caucasian form with *-nx-* ~ *-nx^w* -, where the focus was certainly on the velar consonant. Elsewhere we find consonants that are a subset of those with which *h-* was matched: *-Ø-*, *-h-*, *-x-*, *-γ-*, and *-ʔ-*.

There are another 16 sets involving 38 pairs where *-h-* occurs but where an **-n-* was probably not present [7. 11. 13. 54. 57. 69. 70. 158. 159. 187. 231. 233. 254. 280. 292. 313.] (for *behatz* 'toe, finger' [134.] I accept the inclusion of *hatz* 'finger', and thus count this *h* as initial). A variant with *-g-* instead of *-h-* occurs in 12 of these words. It is time to take off the blinders and recognize that in this *-g-* we are seeing the original consonant, whose loss in some dialects created a hiatus, into which the *-h-* was introduced in Northern dialects.

Contrary to the impression given by Trask, this *-g-* does not generally occur in words for which there is evidence of **-n-*. The only word in our sample showing both is *ziho* (~ *zigo*) 'tallow', for which I have given a tentative explanation involving dual sources [my comment on 305.]. Evidence for a **-g-* is given by the etymology in cases like *uhain* 'wave (n.)' [280. 159.], with **-gain*; I have advocated similar analyses for *zuhain* 'hay, fodder; tree' [313.], also with **-gain*, and *bihar* 'tomorrow' [69.], with **-gar*.

For two of these words showing a variant with *-g-*, a third variant with *-b-* also occurs: *aho* 'mouth' [11.] and *ohe* 'bed' [231.]. This represents the same fluctuation contiguous to the rounded vowel *o* that Trask calls attention to contiguous to *u* in *agure* 'old man' [8.], *sagu* 'mouse' [253.], and *suge* 'snake' [265.]; cf. also my suggestion for *zigo* (< **-b-*) 'tallow' [305.] Yet another example would be *abuztu* ~ *abozto* 'August', from Spanish *agosto* (Gavel 1921:307). It seems that the original consonant might have been either **-b-* or **-g-* in such cases. Another three of these words show a variant with *-r-* [54. 159. 254.]; a fourth has *-l-*, thought to be from **-r-* [57.]; for two of these this is in addition to a variant with *-g-*. The surrounding vowels are different from those in the other words, either being identical, *i_i* and *a_a*, or close to the latter, *e_a*; this may just be an accident of the small sample. One thinks of an original **-d-* here, which is known to have become *-r-* in many words (Michelena 1977:227-29). Thus we see that the source of *-h-* in this case was a voiced stop, predominantly **-g-*, sometimes **-d-*, probably **-b-*, and perhaps also **-r-*. Voiced stops have fricative allophones intervocalically, much as in Spanish, and are recognized as having been subject to sporadic loss. This would be an older stage of development of *-h-* than from **-n-*, since borrowings from Latin do not, to my knowledge, show *-h-* reflecting such consonants.

Interesting additional support for this source of *-h-* comes from the combining forms of certain nouns, contingent upon the loss of *-i* (Michelena 1977:222, 237-38, 249-51). When the preceding consonant is *-g-* or *-d-*, this form ends in *t*: *begi*, *bet-* 'eye' [51.], *idi*, *it-* 'ox'. This same final *t* occurs for stems containing *-h-*: *behi*, *bet-* 'cow', *bihi* ~ *bigi*, *bit-* 'grain, seed', *zohi* ~ *zogi*, *zot-* 'clod', thus implying, of course, that they earlier contained **-g-* or **-d-*. But for stems with *-h-* < **-n-* we find *-n* instead: *mihi*, *min-* 'tongue' [218.].

Trask of course regards the *-g-* as secondary, speaking of "the common western insertion of *g* to separate vowels in hiatus" [7. 69.]. But note that for *ihel* 'frog' [158.] he comments that the variant with *-g-* is "by far the most widespread (and earliest attested)"; furthermore the variant with *-g-* seems to have a definitely eastern, rather than western, distribution in words such as *ahur* 'palm' [13.], *bihotz* 'heart' [70.], and *zuhain* 'hay, fodder; tree' [313.].

In the comparisons made, an *-h-* that may be from **-d-* is matched to Caucasian *l-* because of the variant with *-l-* in *beharri* 'ear' [57.]. For *oihan* 'forest' [233.], there is a matching to three forms with labiodentals *f* and *v*. A *j* occurs twice [231. 254.]. Otherwise the paired consonants are similar to those already noted for *h*: *Ø*, *h*, *k*, *k^w*, *k'^{w}*, *x*, *γ*, *q*, *GG^w*, *q'q'^{w}*, *R^w*, *ʕ^w*, *ʔ^w*. We are hardly seeing regular correspondences here, especially when we remember that this *-h-* may come from three or four different consonants!

The occurrence of aspiration with voiceless stops *p^h*, *t^h*, *k^h*, is a question of stop series, as already discussed, but which enter into patterns of dissimilation with *h* of whatever origin, quite as in Greek, where *h-* < **s-* interacts in a dissimilative pattern with that language's *p^h*, *t^h*, *k^h*. It is intriguing to speculate whether the sequences

of resonant + *h*: *nh*, *lh*, *rh*, might reflect older voiceless resonants (like the sounds spelled *ll* and *rh* in Welsh), but it seems more likely that they originate as, and synchronically continue to be, clusters of consonants. For example, for *anhoa*, Zuberoan *anhúa* 'provisions' from Latin *annōna*, perhaps there was an intermediate form **anoha*, in which *-n-* yielded **-h-* farther along in the word than where it now survives, which then migrated forward in the word; if so, a parallel development in native words for which we have no other evidence of former **-n-* or **-g-* between the second and third syllables might have obtained.

I must comment also on the dozen words from Latin which Trask presents in a scrambled collection (§2[7]), in order to convince the reader that the occurrence of aspiration is hopelessly perverse. There is no doubt that secondary addition and displacement of aspiration has historically occurred, perhaps correlated with accentual differences, as certain shapes of words have come to be favored (Michelena 1977:214-19). But it seems that a good rule of thumb for getting at the earlier form of a word with respect to aspiration would be to look to the aspirating dialects and take from them the form lacking aspiration. Once we do this, we find perfectly well attested forms without an added initial *h*: *are* 'sand', *autatu* 'choose', *arma* 'weapon', *aizkora* 'ax' (but the *h-* in the last might alternatively be due to the presence of *haitz-* 'rock'; cf. my comment on [127.]). Similarly, there are well-established forms without aspiration of the voiceless stops: *piper* 'pepper', *makila* 'stick', *gertu* 'certain', or *-rh-*: *soro* 'field'. In two more, the *-h-* is the regular outcome of *-n-*: *ohore* 'honor', *ahate* 'duck'; in *lehoin* 'lion', it breaks up the hiatus already present in Latin; for two of these, there are attested Northern forms with vowels in hiatus, *oore* and *leoin*; for the third, a form *aate* is attested at least in Western (non-aspirating) dialects. Thus, the only word among these twelve that really presents a historical problem is *anhoa*, discussed in the preceding paragraph.

Finally, I must object to the suggested "Proto-Basque" reconstruction **a-x*^{w}*o* (but according to Bengston's comment **ax*^{w}*o* ~ **axo*) that is proferred for *aho* 'mouth' [11.], as symptomatic of the superficiality with which many of these comparisons have been compiled. *Proto-* (as contrasted with *pre-*) has a specific technical meaning, namely, pertaining to forms that might have occurred in a parent language as determined from the comparison of descendent languages or dialects. Traditionally, the proto-form here has been taken to be **aho*, with loss of the *-h-* in the dialects of Spain and other secondary developments. It will be realized that I am recommending instead **ago*, which survived as such in a broad range of dialects: Bizkaian, Gipuzkoan, High Navarrese, Roncalese, and even some Lapurdian, Zuberoan, and older Low Navarrese. In restricted Bizkaian subdialects, the influence of the rounded *o* led to *abo*. Loss of the medial consonant gave the widespread form *ao* in Southern dialects: Bizkaian, Gipuzkoan, High Navarrese, and Roncalese, this becoming monosyllabic *au* in some Bizkaian and Gipuzkoan localities. A form like this with vowel hiatus led to the general *aho* of Northern dialects: Lapurdian, Low Navarrese, and Zuberoan, while in some Zuberoan, regular raising of the final vowel led to *ahu*. Hence, there is no justification in Proto-Basque for an unneeded additional phoneme **x*^{w} (or two, with **x*).

3. Assumed Affixes

In his §4, Trask also raises the important question of the segmentation of the Basque words under comparison. Certainly the convention should be observed of indicating morpheme boundaries, by hyphens or by extraction of separate pieces, only where this is justified by synchronic analysis of a language, not excluding conservative internal reconstruction. To project morpheme boundaries from one language onto another represents circular reasoning, assuming what needs to be proved, i.e., that the languages are related and/or that the forms in question are indeed cognate. It would also collapse together two different situations, those in which the actual boundaries of compared forms coincide, and those in which they do not. This is not to be taken to suggest that comparisons may not involve only parts of larger forms, as it is perfectly normal for this to be the case; what is wanted is just honesty in reporting. In these data, there are numerous untenable comparisons involving mis-segmentation of Basque words, as will be inventoried in §6.

Trask usefully discusses here several "noun-class" prefixes that are recurrently segmented in Basque words entering into comparisons, and rightly objects to their lack of synchronic justification. Of these, only *b(e)-* on body parts has been commonly thought of previously (see Agud and Tovar [1988-] under **b-** for earlier literature). Here is my own entry into this sweepstakes, based on contemplating the examples conveniently assembled by Trask.

I suggest that the only shape that is historically and semantically justifiable within Basque is *be-* (not *b-* or *bi-*), which would indicate that the body part in question has bilateral symmetry. This would come from the full form of 'two', **biga-*, as the first member of old compounds, with the *-g-* having been lost and the vowels contracting: **bia- > be-*. This preposed shape contrasts with the other reduced form *bi*, which is thought to have arisen in the postposed position (cf. [66.]); it is true that the contraction **-ia- > -e-* is not usually thought of, but its occurrence as a possible sequence was simply not allowed for. We see then that the body parts with *be-* have bilateral symmetry: *begi* 'eye', *belarri* 'ear', *beso* 'arm', and *belaun* 'knee'. Trask is probably right in finding **behe-*, contracted to *be-*, 'lower, below' in *behatz* 'toe, finger' [134.], but note that toes and fingers also are paired off. Given that *beko* 'forehead' [55.] is probably a Latin borrowing, and that my comment on [54.] explains away *behazun* 'bile' as a separate item, we have no counter-examples starting with *be-*. The bilateral symmetry in this context means two saliently separate paired-off items, so that a *b-* or *bi-* in *bihotz* 'heart', *birika* 'lung', *barrabil* 'testicle', or *bular* 'breast, chest' would not convey this meaning. There remain five items starting with *b-* for which this meaning is clearly lacking: *buru* 'head', *bizar* 'beard', *bare* 'spleen', *bizkar* 'back', and *mihi* 'tongue', if showing *m- < *b-*.

We can next check for body parts with bilateral symmetry not bearing this prefix. Trask should not have listed *zango* 'leg' as a counter-example, since, like *hanka* 'leg', it is a probable borrowing from Romance (cf. §2 and [98.]). We are then left with three words: *esku* 'hand', *ukondo* 'elbow', and *oin* 'foot'. 'Elbow' [283.] is a secondary compound, but the other two are certainly old in the language; one therefore infers that 'hand' and 'foot', being at the extremities of the limbs, would be placed so often in disparate positions that they would not count as symmetrically opposed. (In making this suggestion, I have also been mindful of the Wakashan languages of the Northwest Coast, which have lexical suffixes for the most commonly mentioned body parts, of which those naming body parts with bilateral symmetry require *CV-* reduplication of the stems to which they are affixed.)

I regard this suggestion as plausible, but speculative. It of course has nothing to do with Caucasian prefixes. If we now look at the segmentations given in these comparisons, it is a little more favorable to them than the earlier hypothesis, which was less constrained in both form (any shape starting with *b-*) and meaning (any body part). It excuses the lack of segmentation of such a prefix in *beko* 'forehead' (if native), *bare* 'spleen', *barrabil* 'testicle', *bizar* 'beard', *bular* 'breast, chest', and *buru* 'head', and it agrees with the recognition of *be-* in *belarri* 'ear', *belaun* 'knee', *behatz* 'toe, finger', and *behazun* 'bile'. Note moreover that this might offer an explanation for the occurrence of intervocalic *-l-*, rather than *-r-*, in *belarri* and *belaun*, a problem noted by Trask for the former [57.]. There might have been a period when forms with and without *be-* both occurred, much as we see for *hatz* and *behatz* 'finger, toe' in some dialects. Perhaps at one time, e.g., **larri* meant '(an) ear' and *belarri* meant 'pair of ears, both ears'. Then the initial *l-* could have been retained by analogy in the prefixed form. In violation of this hypothesis, we still have the failure to segment *be-* in *begi* 'eye' and *beso* 'arm' (noted by Trask [51. 64.]), and the recognition of *be-* in *bekoki* 'forehead', and of *bi-* in *bihotz* 'heart', *birika* 'lung', *bizkar* 'back', and especially *bizi* 'alive'.

One also finds recurrent segmentation of a few suffix shapes, *-r* and *-ar*, *-el* and *-il*, which are not systematically discussed by Trask and may therefore be summarized here.

There are several sets wherein a suffix **-r* was recognized by Bengtson without synchronic justification, merely because it is not matched in the other forms being compared. These are: [69.] **bihar** 'tomorrow', [75. 288.] **bizkar** 'back', [77.] **bular** 'breast, chest', [87.] **elur** 'snow', and [228.] **nigar** 'tear' (n.). The ending is also segmented as *-ar*, unaccountably in [297.] **zamar** 'sheepskin, sheepskin jacket' and [304.] **zigar** 'mite', and more understandably in [166.] **intzigar** 'hard frost', [294.] **zakar** 'strong, vigorous, brave', and [308.] **zilar** 'silver'. Some synchronic justification, albeit superficial, can be seen for its further recognition in the numerals [128.] **hamar** 'ten' and [206.] **laur** 'four', and the interrogative [302.] **zer** 'what', as well as in [180.] **izter** 'thigh', and might have been applied in [107.] to the comparison of *ganga* 'palate' and *gangar* 'uvula'. Rather obviously, no attempt at suggesting a meaning for this piece seems to have been made. There are another 27 stems ending in *r* which is not segmented (for three of them [4. 7. 117.] there is no *-r* in Čirikba's sets).

In [251.] **sabel** 'stomach, belly', a synchronically unjustified segmentation *sab-el* is shown. This presumed suffix recurs in [217.] *mat-el* 'cheek', where it at least corresponds to a suggested, albeit unlikely, internal comparison. But why not also segment [112.] *gib-el* 'liver', so as to recognize a "body-part" sufix *-el*. More

specifically, since *matel* is probably a borrowing, this could mean 'abdominal organ'. In [281.] a similar *-il* is marked off in *u-kab-il* 'fist', where the ending should certainly be *-bil.*

4. Marginalia

I list here specific comments on many of the numbered comparative sets. These offer very miscellaneous information, but I would point especially to new etymological suggestions for *behazun* 'bile' [54.], *gurpil* 'wheel' [125.], *haitz* 'rock, crag' [127.], *ohoin* 'thief' [232.], *pataxa* 'bottle' [243.], and *zuhain* 'hay, fodder; tree' [313.]; a different reason for rejecting **ika* 'one' [161.]; older forms of pronouns [80. 227.]; and a previously unrecognized sound change [266.].

[1.] Bq **-a** (article). By gradually turning an original distal demonstrative into a definite article, Basque has entered into the Western European area along with Romance, Celtic, and Germanic languages now having this category (much as Hungarian, uniquely among the Uralic languages, also has). The Proto-Abkhaz-Adyghe form is presented as only a demonstrative pronoun, although it has in fact become a definite article in some daughter languages, but since definite articles seem always to have a comparable origin, this latter-day history of Basque should not constitute a counter-argument to this comparison, which however belongs with the material of [135.].

[2.] Bq **aberats** 'rich'. A language conserving this Latin infinitive form only as a noun is Rumanian: *avere* 'property, wealth, fortune'. Note that an alternative trajectory could assume that something like this was the original meaning in earlier Basque, with the meaning 'animal' not having to intervene. The alternation of prevocalic *-r-* in *abere* with preconsonantal *-l-* in examples such as *abelbide* 'trail for animals', *abelburu* 'head of cattle' (pointed to by Bengston in his comment) does not disprove that this was an original *-r-*, since a synchronic pattern has established itself of stem-medial *-r-* becoming stem-final *-l-* before added elements. Note, on the one hand, examples like *abelaska* 'manger', *abeletxe* 'stable, stall', where the *-l-* occurs also before a vowel; and on the other, ones like *amoltsu* 'affectionate, kind, amiable', from *amore* 'love' < Latin *amōre* 'id.', where there is clear evidence of an original *-r-* (cf. Michelena 1977:317-19).

[5.] Bq **adin** 'age'. **Aditu** is discussed below, not above.

[7.] Bq **agor** 'dry'. There are also transitional eastern forms *ihar* and *lehor.* And just so the situation doesn't seem too simple, one should remember a fourth widespread word *idor.*

[8.] Bq **agure** 'old man'. In support of the derivation from **avule*, see Corominas 1969:177-82. The postulated derivative of Latin *avu* 'grandfather', **avulu*, although not attested can be inferred from forms like Italian *ávolo* 'grandfather'. Corominas points to other reflexes of vocative forms, such as Basque *done* 'saint' < Vulgar Latin **domne* < *domine*, vocative of *dominu* 'master'. Corominas goes on more controversially to relate this word to *guraso* 'parent' < **aguratso*, in competition with a derivation from *gure* 'our' (Tovar 1959:32-7).

[9.] Bq **ahal** 'ability'. It can be explained that the *-ge*, *-ke* in *ahalge* 'shame' is a reduced form of *-gabe* 'without, -less'.

[11.] Bq **aho** 'mouth'. See the end of §2 for information on dialect variants and a criticism of the "Proto-Basque" reconstruction **a-x^{w}o.*

[12.] Bq **ahuñe** 'kid'. The evidence for **anu(n)-* consists only of the plausible equating with *a(h)untz* 'female goat' of the Bizkaian family name and toponym attested as *Anuncibay*, now *Auntzibay*, where the second part would be *ibai* 'river', hence literally 'Goat-river'. (Here and in [159.] and [305.] the reference should be to Michelena 1950c rather than 1950a.)

[23.] Bq ****-ar** (plural suffix). I have wondered whether the absolutive plural *-it-* (as in *ditu* 'he has them' vs. *du* 'he has it') might be equated with the *-ir-* in *dira* 'they are' (vs. *da*), both of them applying plurality to a preceding *d-* [80.], but find no parallel for a *-t-* ~ *-r-* alternation. Such an **-ir-* would be the best approximation to the assumed **-ar*. Castaños Garay (1979) is a whole monograph devoted to Basque plural formation.

[25.] Bq **ardo** 'wine'. Additional evidence for reconstructing an **-n-* here is found in the Lapurdian and High and Low Navarrese dialect form *arno* and the Low Navarrese form *ano*.

[30.] Bq **arroda** 'wheel'. In further support of this word's being a Romance borrowing are the facts that it is limited to the Lapurdian dialect, and that geographically closer forms are a little more similar: Gascon *arroda* and Béarnais *arrode*. It might be mentioned that the earlier borrowing *errota* commonly has the meaning 'mill' and only in Zuberoan that of 'wheel'.

[32.] Bq **astigar** 'maple'. Bengtson certainly applies a double standard in his handling of this word, on the one hand criticizing Trask for preferring a Basque form of the central dialects (namely Lapurdian and Low Navarrese), on the other comparing to a word in a single North-Central Caucasian language, Batsbi, which gives it no pedigree in the Caucasian family. The resemblance is moreover too pat to be plausible.

[40.] Bq **azkoin** 'badger'. Since the comparison with the first-listed form *azkonar* is to the first part, equated with *hartz* 'bear', this pair should be merged with [132.].

[46.] Bq **bargo** 'young pig (3-6 months)'. Why not mention the Latin masculine form *porcu* 'pig, hog' (Spanish *puerco*), whose final vowel then matches. Similarly, note the widespread Spanish variant form *varraco* 'boar'.

[48.] Bq **bartz** 'louse, nit'. The reader could be helped by being told that the information in Corominas and Pascual (1980) concerning 'louse' is found under the word for 'pimple', Castilian *barro*.

[50.] Sp **becerro** 'bullock'. This word is at least indigenous to the Iberian Peninsula, having a cognate only in Portuguese *bezerro* 'id.', and is thought by Corominas (1954[1]:434-5) to be connected with *bicerra* 'mountain goat' and to bear an Iberian suffix **-err* or **-irr*.

[51.] Bq **begi** 'eye'. For *bepuru* 'eyebrow' the reader should also be referred to [78.] **buru** and my comment thereon.

[52.] Bq **beginini(ko)** 'pupil (of the eye)'. It should be made clear that Spanish *niña* also means 'girl', and the parallel Spanish expression *niña del ojo* 'pupil', literally 'girl of the eye', should be pointed to, with the same allusion to the small images seen in the pupil as underlies English *pupil*, ultimately from Latin *pūpilla* 'girl, doll, pupil', in turn influenced by, or at least paralleling, Greek *kórē* (κόρη) 'id.'

[54.] Bq **behazun** 'bile'. The observations that this reconstructed stem looks like a word for 'finger' [134.], and that at least one form means also 'gall bladder' (Agud and Tovar [1988-] indicate general meanings 'bile' and 'gall bladder of fish'), lead to my etymological suggestion: the original meaning was 'gall bladder', and the stem for 'finger' is indeed present, so that the literal meaning was 'having a finger' and perhaps also 'place of finger'. The gall bladder has a duct that might be likened to a finger. This suggestion, pointing to a stem *hatz* 'finger', would obviously ruin this comparison.

[55.] Bq **beko** ~ **moko** 'forehead'; 'beak, extremity'. The two words *beko* and *moko* should be kept apart. Corominas (1954[3]:404-06) gives primacy for Spanish *mogote* to a derivation from a presumed Basque **mokoti* 'sharp pointed' derived from *moko*. *Beko* is certainly from Latin *beccu* (along with French *bec* and Spanish *pico*), to which however is commonly imputed a Celtic origin.

[56.] Bq **bekoki** 'forehead'. As per my comments on [51.] and [78.], in relationship to Trask's attractive suggestion there would have been an early form **bet-goi-ki*.

[57.] Bq **belarri** 'ear'. My suggestion above (§3) might justify a prefix *be-* here, as well as the occurrence of intervocalic *-l-*.

[58.] Bq **belaun** 'knee'. This word, like *belarri* 'ear' [57.], offers the problem of occurrence of intervocalic *-l-* rather than *-r-*. Again, my suggestion above (§3) might explain this and justify the recognition of *be-*.

[59.] Bq **beltz** 'black'. *Beles* (and *-beleś* in indigenous script) is Iberian rather than Aquitanian; the latter shows *belex*, in which, as noted in [70.], the *-x* represents an affricate (cf. Michelena 1964a:17; 1977:63, 162, 327, 416). I don't see why assumed **beletz* should constitute a defect of the comparison, since the Proto-North-Caucasian form also has two vowels. Regarding *harbel* 'slate', one even notes a name *Harbeles* in an inscription (CIL 13.1.85), which of course need not be equivalent.

[68.] Bq **biga** 'two-old-year heifer'. An intervocalic **-n-* and thus the Latin etymon **bimāna* is confirmed by the nasalization and stress placement in Zuberoan *bigã́*, and probably by the *m-* in High and Low Navarrese *miga*.

[69.] Bq **bihar**. In spite of the meaning 'yesterday' carelessly given in the lemma, this word does mean 'tomorrow'. I definitely endorse the derivation from 'second'. Ordinal numerals are formed with the suffix *-garren*, so 'second', formed from *bi* 'two', is *bigarren*. But this must be a locative formation with *-en*, hence 'tomorrow' would be from **bi-gar* (or **biga-gar*). Note the parallel derivation of *bakar* 'only, unique, alone' from **bat-gar* 'first'. This suggestion is favorable to these comparisons in justifying a **-g-*, but not otherwise.

[70.] Bq **bihotz** 'heart'. A variant form *bigotz* occurs in High Navarrese and Roncalese, which is relevant to my discussion of the origin of *-h-*, §2. Whether the Aquitanian names with *-h-* are this same word is, of course, uncertain. But there is no matching consonant in either form compared.

[73.] Bq **bizar** 'beard'. My suggestion above (§3) might justify the lack of segmentation of *b-* here.

[74.] Bq **bizi** 'alive'. The reader may wonder why a stem *b/pitz-* is mentioned. This occurs prevocalically in forms not shown and would be an analogical formation from *b/piztu* based on the fact that the preponderance of stems that can occur in word-final position end in an affricate (*-tz*, *-ts*, *-tx*) rather than a fricative (*-z*, *-s*, *-x*), and that before a consonant affricates are replaced by the corresponding fricatives. The final affrication has even occurred in borrowings such as *gorputz* 'body' from Latin *corpus*. Thus, from *bizi* would have come first *b/piztu* by replacement of the suffix *-i* by *-tu*, with *b/pitz-* developing later before vowels after the two stems became semantically separated (cf. the relationship in [139.] between *hatz* 'lineage, race' and *hazi* 'to grow, grow up', and Michelena 1977:288-89).

[75.] Bq **bizkar** 'back', 'height in mountain'. The logical possibility exists that a historically present stem starting with *zk-*, like **-zka(r)*, might occur, having survived after vowel-final prefixes, even though such a word-initial consonant cluster is excluded. I can't find any morphemes beginning with *zk-*, *sk-*, *xk-*, though; either the cluster is morpheme-internal, as in *azken* 'end' and *ezko* 'wax', or a morpheme boundary historically separates the consonants, as in *hizka* 'letter of the alphabet', *harrizko* 'made of stone', and *dizkiot* 'I have them for him'.

[77.] Bq **bular** 'breast, chest'. My suggestion above (§3) might justify the lack of segmentation of *b-* here.

[78.] Bq **buru** 'head'. Regarding *bep(h)uru* 'eyebrow', it might be explained that *begi* 'eye' has the combining form *bet-*, as seen in the words for 'eyelash' and 'eyelid' shown here, so that this would come from **bet-buru*, quite parallel phonologically to its synonym *bekain* < **bet-gain*, with *-gain* 'over, above'. My suggestion above (§3) might justify the lack of segmentation of *b-* in *buru*.

[79.] Bq **busti** 'moist, wet'. For Old French *moiste*, the source of English *moist*, a widely favored alternative source is Vulgar Latin **muscidu* 'moldy, wet', a variant of *mūcidu* 'slimy', still probably influenced by *musteu*.

[80.] Bq **d-** (verbal prefix). This prefix uniquely resisted the regular change of **d-* to *n-* [my comment on 266.]. This may perhaps be due to its frequent occurrence after preverbal particles, especially affirmative *ba-*, e.g., *badakit* beside *dakit* 'I know it', giving transparent analogical support from the internal position in which the *-d-* was preserved; also, in contrast to first person singular *n-* beside forms of independent *ni*, to the absence of a clear relationship to an independent pronoun. In Jacobsen 1975, I suggested its origin in a pronoun **do*, pieced together from the alternants *d- ~ -o- ~ -Ø*, much as second person singular **ga* is inferred from *h- ~ -a- ~ -k*. Remembering the change to *n-*, can we be seeing this pronoun in the interrogative stem *no-* [229.]?

[89.] Bq **emazte** 'wife' The same objection to the shape of a presumed stem **-zte* would apply as to **-zka(r)* in *bizkar* 'back'; see [75.] and my comment thereto. The derivation from **emagazte* 'young female' is bound to be correct and scarcely needs justification. It's true that we expect more 'old', as in colloquial American English *old lady* and *old woman* for 'wife'. But *young lady* can mean 'a man's sweetheart or girl friend', and people got married younger in those days!

[90.] Bq **eme** 'sweet'. There is no such Basque word. We probably have to do with *eme* 'female, feminine', discussed under [88.], which has extended meanings 'soft, gentle, smooth, slow'. Thus this is 'sweet' much as in *she's sweet*, but not *it's sweet*. A mistaking of a French gloss 'doux' or a Spanish one 'dulce' probably occurred, which mean both 'gentle, soft' and 'sweet', much as Trask notes for a Spanish gloss in [238.]. This correction certainly invalidates this comparison to Caucasian.

[94.] Bq **eri** 'sick'. Although not reproduced in the lemma, this word indeed also occurs as a noun meaning 'sickness'.

[98.] Bq **esne** 'milk'. There is a widespread variant form *ezne*; it is unclear which one is more archaic.

[100.] Bq **etxe** 'house'. Although *etxola* 'hut, cabin' is presented segmented as *e-txo-la*, it seems likely that the last syllable is intended to match the **-rHV* of Proto-North-Caucasian **c'VrHV* 'house', which is why it is also cited. However, this is just a compound of *etxe* with *ola*, itself meaning 'hut, cabin'.

[107.] Bq **ganga** 'palate'. Spanish *canga* is thought by Corominas (1954[1]:632-3) to be derived from a Celtic word for 'curved wood' and hence to be related to *cama* in certain meanings.

[109.] Bq **garkotx(e)** 'nape'. See also [198.] for *kokot(e)*. The meaning of *kokots ~ kokotz* is predominantly 'chin' rather than 'nape'.

[120.] Bq **gose** 'hungry'. Although not reproduced in the lemma, this word indeed does occur as a noun meaning 'hunger' (like the Caucasian form compared).

[122.] Bq **gu** 'we'. As mentioned in [227.], this is also a suffix *-gu*, as in *dugu* 'we have it', and it has a shortened prefixal form *g-*, as in *gara* 'we are'.

[125.] Bq **gurpil** 'wheel', **inguru** 'vicinity', 'around, near'. It seems to me that the *gur-* of *gurdi* 'cart', and hence of *gurpil* 'wheel', might itself be borrowed from Latin *gȳru*, which means 'ring' as well as 'circle' (whence also Spanish *giro* 'rotation, turn, etc.' The suffix *-di* can indicate a 'collection of', so that 'cart' would have been literally a 'collection of wheels'. (In [Hokan] Washo, the word for 'hoop', *pú·lil*, was extended to mean 'wheel' and 'wagon', now usually 'car'.) The competing Latin etymon has been *curru* 'chariot'. It might be explained that *gurdi* has a combining form *gurt-*, as in *gurtabere* 'animal used for pulling a cart', so that there would have been an

intermediate form **gurt-bil* behind *gurpil* (cf. my comment on [78.]). Perhaps we see a direct combination of the first and last parts in *gurbil* 'small barrel'.

[127.] Bq **haitz** 'rock, crag'. The evidence for **-n-* comes from the Roncalese *aĩnz-* occurring in the words for several implements that have often, but controversially, been thought to be derivatives of this word, and therefore remnants of stone age nomenclature, such as *aĩnzto* 'knife' and *aĩnztur* 'shearing scissors'; elsewhere, we find forms with *(h)aiz-* and *(h)aix-*, such as *aizto* and *aiztur*. Also, Eastern forms for 'hoe' with *-n-*, *(h)aintzur*, *antxur*, also *altxur*, are noted, which elsewhere is *(h)aitzur*. Michelena (1977:206) also lists several medieval names which document the *h-* but not the **-n-*. The fact that the intervocalic **-n-* here did not become *-h-* points to an initial *h-* in the original form, by the Grassmann's-Law-like principle that there can be only one aspiration per word, hence **hanitz* or **hanetz*. Initial *h-* may come from **k-*, which we perhaps see in *harkaitz* 'rock, crag', where the *har-* is a combining form of *harri* 'stone, rock' [131.], also seen in *harbel* 'slate' [59.]. Trask himself suggests an earlier **k-* in *harri*, and we might further suppose that *har-* and **han-* are alternate forms of the same morpheme, perhaps from **kar-* ~ **kan-*, showing the *-r* ~ *-n* alternation pointed to by Trask in [83.]. Yet more speculatively, the ending of *haitz* might be *(h)a(i)nitz* 'many' (relating to the meaning 'crag'), hence < **hanhainitz*.

[128.] Bq **hamar** 'ten'. If we accept the comparison to Sumerian *haw(a/u)mu*, its *-w-* could attest to another possible source of *-m-*: by contamination from the following nasal *-m-*, quite as in *hemen* (beside *heben*) 'here', from a locative formation in *-en* on the proximate demonstrative *hau*, **hawen* giving **hewen* by metaphony, giving *hemen* by nasal assimilation (Michelena 1977:177, 275); yet another example is seen in Trask's discussion under [249.]: *eraman* 'to take away' < **erawan* < **e-ra-oa-n*. The compound of *hamar* alluded to is *amabi* '12', giving **amabitu* > *amaitu* 'to finish' (i.e., after handling a dozen you're finished) (also Michelena 1977:496), which is worth mentioning because it shows the same dissimilative loss of *-b-* as in the form that probably underlies *hamaika* '11' (see my comment on **ika* [161.]).

[130.] Bq **haragi** 'meat, flesh'. It's probably worth making clear to those readers who may be consulting separate entries out of order that this discounting of *h-* for etymological purposes is what Trask is in fact recommending. But note that, as mentioned above, there are some 29 other comparative pairs wherein Basque *h-* is not matched by any consonant in the other language.

[131.] Bq **harri** 'stone, rock'. In support of original **k-* here, see my comment on [127.].

[135.] Bq **hau** 'this', **han** '(over) there', **hara** 'thither'. Additional support for the possibility that Basque initial *h-* might sometimes come from **k-* (cf. [131.] and my comments on [127.] and [145.]) comes from the fact that the demonstrative forms in the easternmost dialects, Zuberoan and Roncalese, begin with *k-* instead of *h-*, and those in adjacent Aezkoan and Southern High Navarrese begin with *g-* (Michelena 1977:246-47). This fact, however, is not helpful to these comparisons to Caucasian forms, which lack initial consonants.

[140.] Bq **hegal** 'wing'. My comment on [130.] applies here also. The suggested segmentation *he-gal* is somewhat supported by the Roncalese form *magal*. *Hegal* and *egal* are forms of the eastern dialects; the corresponding western form is *ego*. If this implies a segmentation *heg-al*, it would greatly weaken the comparison. Yet another segmentation is suggested by the comparison to *hegatz* ~ *hegatx* 'feather'.

[145.] Bq **hi** 'you, thou' (singular intimate). Back in the days when *zu* 'you' (singular polite) was plural [312.], *hi* was the corresponding singular, with no distinction of intimacy. There are no prefixed forms *hi-* ~ *he-*; these were probably taken by Bengston from the first part of words in which it is the stem, such as *hire* ~ *here* 'your'. The comparison with the suffixed form *-k* (inappropiately repeated as [191.]) gives a reason to suppose that *h-* here might indeed be etymological, i.e., it might come from **k-*, a possibility that Trask entertains elsewhere ([131.]), or else from **g-*.

[151.] Bq **horma** 'ice'; 'wall'. It's true that Latin *forma* had neither meaning. For the latter, Michelena (1964b:132) draws attention to its Gascon reflex *arroume* 'wall of stones piled up without mortar'. In Castilian, there is *horma* or *pared horma* with the same meaning. Regarding the former, the early attestation gives the meaning 'coating of ice' (French 'verglas'), still reported in some sources, i.e., ice that has "formed". I would not assume that we have "insertion of unetymological *h*-" here, but rather that the *h*- was a direct substitute for Latin *f*- in certain dialects.

[159.] Bq **ihintz** 'dew'. I find the reconstruction **inintz* unlikely and think that a good case can be made for **iCintz*, as an alternative to **initz*, where *-*C*- is a non-nasal, probably *-*d*-. In Bizkayan, we find *iruntz* and *irauntz* (Michelena 1977:80-81). We don't need more than one nasal consonant in order to account for the nasalization in Zuberoan *íhĩtz*. Another Bizkayan form *iñontz* points to nasal infection of a glide *-*y*-, rather than irregular retention of *-*n*-.

[161.] Bq ****ika** 'one'. It is scarcely fair to characterize the widely shared idea that *-*ika* (or *-*eka*) might mean '1' as a "wild guess". This is a straightforward extrapolation from the pattern for forming the numerals '12' through '18', which compound *hama*- '10' (*heme*- in '18') with the lower numeral representing the additive amount, e.g., *hamazazpi* '17' (*zazpi* '7'). Some have not been able to resist the comparison to Sanskrit *éka*- '1', where, however, the *e*- is structurally a diphthong (< **ai-ka*- < **oi-ko*-). Trask's suggestion of a different original meaning for this piece is attractive (cf. the fact that the English additive formation for the teens doesn't get started until '13'). I much prefer, however, Michelena's (1977:496) suggested reconstruction **hama-bed-ka* (better **hama-bet-ka*), wherein the middle piece is a form of *bat* '1', and *-ka* can mean 'in groups of'. Thus, I am in agreement in rejecting this presumed word, but on the basis of form rather than meaning.

[163.] Bq ****ilu**. Of course, a contemporary **ilu* could have come from **illu* (Michelena's **iLu*) (cf. [123.]).

[170.] Bq **itain** 'tick'. Bq *tiña* 'ringworm' should be added to the lemma and cross-referenced in alphabetical order.

[181.] Bq **izten** 'awl'. The predominantly occurring form is *ezten*.

[183.] Bq **izu** 'trembling', 'fear'. The eastern dialects show a less similar variant *izi*, but this is probably due to secondary assimilation.

[185.] Bq **jakin** 'know (a fact)'. It looks as though the *j*- here is being equated to Proto-North-Caucasian **H*- and Burushaski *-h*-. But as in nearby items, this is from a vowel **e*-, thus **e-aki-n*.

[187.] Bq **joan** 'go'. It is not clear whether the *j*- is being equated to vowels, Proto-East-Caucasian *-*A*- or Proto-Yeniseian *-*e*-, or consonants, PEC **ʔ*- and PY **h*-. But, again, this is from a verbal **e*- (via **y*-) (cf. forms such as *noa* 'I am going'). It will be seen that this stem shows a hiatus of vowels *-oa*-. In the Southern *h*-less dialects, this has been resolved by a loss of syllabicity, **yo*- > **yw*-, with varied outcomes in tandem with the outcomes of **y*- itself: *xwan, šwan, fuan, fan, gan* (Michelena 1977:173). In some Northern *h*-full dialects, on the other hand, there occur synthetic forms with *-h*-: *noha* 'I am going', etc. There is also a form *johan* used like the present participle *joaiten*, allowing a distinction between *johan da* 'he is going' and *joan da* 'he has gone' (Lhande 1926, s.v.). These facts lead me to reconstruct a medial consonant: **e-oCa-n*.

[189.] Bq **-k** (plural suffix). I think the plural suffix was just *-*g* (word-final stops have been unvoiced).

[191.] Bq **-k** (second person singular [intimate male] agreement marker in verbs). This set should be merged with [145.].

[193.] Bq **kako** 'hook'. Castilian (and Portuguese) *gancho* 'hook' is probably a pre-Roman word, derived by Corominas (1954[1]:656-9) from Celtic **ganskį̣o* 'branch'.

[197.] Bq **karats** 'bitter'. Given its other meaning 'foul-smelling, fetid', one wonders about a connection to *k(h)orotz* 'dung' [199.]. Moreover, *k(h)irats ~ k(h)irets* 'stench, fetidness, rotten smell' must be taken into account.

[198.] Bq **kokot** 'nape'. *Cocote* is attested from the 16th century and so should not be labeled "Old" Spanish. Suggestive of borrowing is the limitation of this Basque word to the westernmost dialects in Spain: *kokot* to Bizkaian, and its variant *kokote* [109.] to Gipuzkoan and nearby High Navarrese.

[200.] Bq **kuma** 'mane, horsehair'. In addition to the limitations pointed out by Bengston: that the Spanish word is not attested before the 17th century (so should not be called "Old" Spanish), is now obsolete, and is of a literary cast, the vocalism is also not right, as a Latin *o* should yield a Basque *o*, and this discrepancy is worsened by the form *k(h)ima*. If, however, the word may have come down by oral transmission into older Castilian, the form **cuema* would have been regular (the Latin vowel being short), which would fit nicely. The form *k(h)ima*, which occurs with this meaning in Lapurdian and High Navarrese, as well as Gipuzkoan, may result from confusion or contamination with another widespread word of this shape (and also *kimu*) meaning 'sprout, shoot, bud, tip of branch', from Latin *cȳma* (whence also Spanish *cima* and French *cime*).

[203.] Bq **lagun** 'companion, friend'. In his comment on Bouda's set referred to by Trask, Kuipers (1960:109) criticizes a comparison to West Circassian *leɣʷə* 'coeval' (in Kuipers' analysis *laǧʷ*) in that this probably contains a suffix *-ǧʷ* 'companion'. The first part of the word is more tentatively compared to a *-la* in 'boy' and to *l* 'flesh'. The Proto-Nakh-Daghestan form compared by Čirikba may conceivably contain a cognate of this suffix.

[206.] Bq **lau** 'four'. I'm afraid that the resemblance seen by Čirikba between Basque *lau* and Proto-Abkhaz-Adyghe **p-kˇ'ə* is beyond my limited powers of imagination.

[216.] Spanish **marrano** 'pig'. An Arabic source is highly likely for this word (Corominas 1954[3]:272-75).

[217.] Bq **matel** 'cheek'. There occur also the forms *mazela* and *masaila*. The variant **maxella* of Classical Latin *maxilla* 'jaw' is so marginally attested that it should be starred. The change of meaning from 'jaw' to 'cheek' is paralleled by Spanish *mejilla* and widely in Western Romance.

[218.] Bq **mihi** 'tongue'. The evidence for the **-n-* is very strong, even if the connection with *mintzo* 'voice, speech, conversation' might be questioned. On the one hand, there is nasalization of vowels in Zuberoan *mĩ́hĩ* and Roncalese *mĩ*; on the other, we see *min* both in isolation in Gipuzkoan and Bizkaian, and as a combining form in Gipuzkoan and High Navarrese *mingain* (compounded with *gain* 'upper surface'). But the comparative evidence in fact points to **mini*, as all of the dialect forms have *m-*, and it is interesting to note that Michelena (1950c:450) earlier gave just this reconstruction. The setting up of **bini* is due to 1) the presumption that there was no **m-* in Proto-Basque, 2) the finding here of a body-part prefix **b-*, whose existence Trask however rejects (§4), and 3) the occurrence of the *-n-*, so that spread of nasalization giving **b- > m-* can be assumed, a process for which there are numerous parallels (e.g., High and Low Navarrese *miga* 'two-year-old heifer' < Latin **bimāna* [68.]). (More pertinent pages in Michelena 1977 would be 268-69 and 275-76. Michelena 1958, cited here and in [289.], is not in Trask's list of references.)

[221.] Bq **mulo** 'pile, stack (of hay or cornshocks)'. In further support of this word's being a borrowing from Gallo-Romance is the fact that it is attested in this meaning only in Lapurdian, a dialect of France.

[223.] Bq **mutur** 'snout, muzzle'. Corominas (1954[3]:389-400) expresses complete uncertainty as to whether this word is indigenous to Basque or borrowed from Romance. He mostly focuses, however, on its meaning 'angry, irritated'.

[227.] Bq **ni** 'I'. There are no prefixed forms *ni-* ~ *ne-*; these were probably taken by Bengston from the first part of words in which it is the stem, such as *nire* ~ *nere* 'my' (beside the archaic form *ene*). See my comment on [266.] for evidence that the forms were originally **di* and prefix **d-*, having undergone the regular change of initial **d-* > *n-* that is the main reason why native words do not begin with *d-*.

[229.] Bq **nor** 'who'. See my comments on [80.] and [266.] for evidence that this form may come from earlier **do-*.

[231.] Bq **ohe** 'bed'. Variant forms *oge* in Bizkaian and Gipuzkoan and *obe* in High Navarrese also occur, which are relevant to my discussion of the origin of *-h-*, §2.

[232.] Bq **ohoin** 'thief'. The recognition of an **-n-* here is based on the nasalized vowel in Zuberoan *ũhúñ* 'thief'; there is also a related Western verb *ostu*, *oostu*, *onstu* 'rob, steal' (Michelena 1950c:450, 1977:303). This gives a reconstructed form **onoin*, which leads to the thought that it may etymologically be a compound of *on* 'good' [235.] + *oin* 'foot' [234.], literally 'good-footed', implying quick at making a getaway. There are parallel compounds such as *onbide* 'good example, virtue', literally 'good way' [67.]. This would of course remove this word as a separate entry.

[239.] Bq **oskol** 'shell, peel, bark (of a tree)'. The Castilian form is *cuscurro*. Given the distribution of the Romance words, a borrowing into them from a Pyrenean substratum seems likely.

[242.] Bq **paru** 'stick of wood, pole'. With *pau*, compare Béarnais *paou*, *pau*. Otherwise I find forms pointing to a Romance source like Spanish *palo*: Southern Lapurdian *paro* and more widespread *palu* and *palo*.

[243.] Bq **pataxa** 'bottle'. The suggested comparison to a word for a 'sailing vessel' inappropriately brings in a different meaning of vessel. I introduce my own suggestion by pointing to the existence of two variant forms, *potaxa* and *botexa* occurring in nearby valleys on the Spanish side of the Pyrenees, respectively Baztán and Salazar, and which mean 'earthen jug' rather than 'bottle'. These words are matched by Aragonese *boteja* 'earthen jug', cognate with French *bouteille* 'bottle', coming from late Latin **butticula*, diminutive of **butte* 'cask' (also the source of English *butt*). French *bouteille* was borrowed as Castilian *botella* and Basque *butila*, *botila*, *botoila* 'bottle', and has probably affected the meaning of Basque *pataxa*. Castilian *botija* 'earthen jug' would come from a variant diminutive **buttīcula*. (in older Castilian and Aragonese, the *-j-* was [š], the sound represented by Basque *-x-*). The source of this Latin word is uncertain; both Celtic and Greek have been pointed to.

[244.] Bq **pinpirin** ~ **pinpilin** 'butterfly'. Influence from French *papillon* 'butterfly' and/or Béarnais *pamparole* 'small butterfly' is also rather likely here.

[245.] Bq **pintza** 'membrane'. A widespread variant form *mintz* occurs, whose *m-* greatly strengthens the assumption that the source word had a *b-*. However, the specific source assumed by Corominas (1954[1]:460-61), Vulgar Latin **vinctiāre* 'to tie', from the past participle *vinctu* of Latin *vincīre* 'to bind, tie around' seems too speculative to be convincing. Thus a borrowing into Pyrenean Romance seems likely here.

[246.] Bq **pipil** 'bud'. There exist additional (Northern) Zuberoan forms, adjective *t(h)ipil* and verb *t(h)ipiltü*, whose initial consonant offers a link in position to the Latin *d-* and in voicing to the Western *p-*.

[251.] Bq **sabel** 'stomach, belly'. The unjustified segmentation *sab-el* was discussed above (§3).

[254.] Bq **sahats** 'willow'. My first guess would be **sagats*.

[257.] Bq **sasi** 'bramble'. Spanish *zarza* 'bramble' goes with Portuguese *sarça*, but otherwise has no clear source. Corominas (1954[4]:853-56) compares it to yet other words presumably indigenous to the Iberian Peninsula.

[259.] Bq **senton** 'old man', **sentana** 'old woman', **santan** 'ever'. *Senton* is doubtless also derived, like *santan*, from French *cent ans* 'a hundred years'.

[266.] Bq **-t** 'I' (ending of verbs). It is an interesting question whether the relationship of this suffix *-da-* ~ *-t* to the prefix *n-*, also seen in the independent pronoun *ni* [227.], is historically one of suppletion or identity. I think the latter is the case, and Martinet (1955:388), for one, expressed tentative agreement. He mentioned initial and final forms *n-* and *-t* < **d* (but not attested *-d-* in *-da-*), and thought of an original complex phoneme $*n^d$. One can also point to the strikingly similar example of a (Siouan) Hidatsa phoneme with initial allophone [n] and medial [r], and for which a now extinct dialect had an additional allophone [d] (Voegelin and Voegelin 1959:15). Martinet did not allow, however, for the absence of native words beginning with *d-* (as noted by Trask §2[1]), so that no additional phoneme need be set up. This would have been **d*, which contrasts intervocalically with **-n-* (> *-h-*), but became *n-* initially by a regular sound change, probably merging with an already present **n-*. Note that by grouping together this prefix and suffix we attain a pattern of vocalism that is shared with the second person singular pronoun: *C-* ~ *-Ca-* ~ *-C*, independent *Ci*, with consonants **d* and **g*, just as there is a shared pattern in the plural pronouns of these same persons: *C-* ~ *-Cu-* ~ *-Cu*, independent *Cu*, with consonants *g* and *z*.

[267.] Bq Bq **-ta** ~ **-eta** (empty morph). I don't agree that *-ta* is "empty" in the sense of 'meaningless'; it rather has a meaning of indefiniteness. It occurs also in singular indefinite forms: *menditan* 'on (a) mountain'. In Jacobsen 1975, I showed that *-eta-* marking certain plural cases comes from **-e-g-e-ta-*, and questioned its derivation from the Latin ending. The *-eta* in place names, on the other hand, may indeed come from Latin. Castaños Garay (1979:66-68) gathers together several older theories about this.

[270.] Bq **txahal** 'calf'. The evidence for former **-n-* here comes only from the nasalization in Roncalese *xãl*, but which is lacking in Zuberoan *xahal*.

[271.] Bq **txar** 'bad'. The derivation of *txar* 'bad' from *zahar* 'old' (itself compared in [292.]) is unquestionable. It is bolstered by attested intermediate forms, *-txahar* 'ugly' and *-zar*, *-txar* 'old, worn out' (Michelena 1964b:96; 1977:191, 413). This certainly makes the proposed comparisons, which were already semantically dubious, untenable in both form and meaning. But note that if accepted, it would give us yet another sound correspondence, Bq *z-* : Cauc. [**č'-*], to be added to the tabulation in §4.

[275.] Bq **txori** 'bird'. This memorable etymological suggestion of Michelena's can hardly be doubted. (Whenever I encounter the common expression *zorionak* 'congratulations, good wishes, etc.', I can't help remembering that it is etymologically 'good birds'.) The PNC **č'w-* should probably be corrected to **č'ʷ-*, but in any case, the substitution of the original form *zori* here will give us yet another sound correspondence for Basque *z-* to be added to the tabulation in §4.

[280.] Bq **uhain** 'wave (n.)'. This analysis is strengthened by the occurrence of related forms showing *-g-*: *ugañ* 'surface of the water', 'otter' (Zuberoan *ügañ* in the latter meaning).

[283.] Bq **ukondo** 'elbow'. I am not troubled (as Bengston is in his comment) by the initial of *ondo*, from Latin *fundu*, since *h-* is a common substitute for Latin *f-*, as in *horma* [151.]; it either did not survive when it began a third or later syllable of a word, as in **ukohondo* or **ukagalhondo*, or may indeed be the aspiration we see in *ukhondo*.

[286.] Bq **urtxakur** 'otter'. There is no piece matching the *ur-* 'water' in the two words compared. Therefore, this set should be merged with [296.] *zakur* as merely comparing 'dog' to yet other mammals (in addition to 'donkey').

[288.] Bq **uzki** 'anus, buttocks'. The same objection to the shape of a presumed stem **-zki* would apply as to **-zka(r)* in *bizkar* 'back'; see [75.] and my comment thereto.

[289.] Bq **xahu** 'clean'. Traces of the *-n-* are seen in the nasalized vowels in Roncalese *xaĩ(n)* and in the derived verb, Zuberoan *xãhãtü*, Roncalese *xaũtu* 'to clean'. How frustrating this set must be. The Proto-Yeniseian **-r- ~ -l-* must have been intended to correspond to the *-n-* in the presumed cluster *-nh-*, but this must be corrected to *-h-*, which, however, comes from **-n-*, but then the word certainly came from Latin!

[292.] Bq **zahar** 'old'. A variant *zagar* is attested in the High Navarrese of the Baztán Valley, which is relevant to my discussion of the origin of *-h-*, §2. It also seems likely to me that the widespread word *zabar* 'lazy, slow, careless, abandoned, corrupt, etc.' is etymologically this same word, although it does not show the rounded vowel elsewhere associated with the *g ~ b* variation.

[294.] Bq **zakar** 'strong, vigorous, brave'. Note that *zango ~ zank(h)o* 'leg' is itself compared in [298.].

[305.] Bq **ziho** 'tallow'. This also means 'fat, grease'. The evidence for **-n-* comes from the nasalization in Roncalese in the derived verb *zĩatu* 'cover oneself with grease; (oil) to coagulate', and, if related, the adjective *zĩ* 'sour, rancid'. Could we also be seeing a combining form *zin-* in *zingar*, *xingar*, *txingar* 'bacon, ham, lard', the last also 'crisp piece of fried pork skin' (where the second part would be *gar* 'flame'), as well as Zuberoan *txinkhor*? Bengston (in his comment) correctly notes the absence of nasalization in Zuberoan *ziho*, but a change from *-h-* to **-n-* need not invalidate his comparison at least to a Proto-North-Caucasian form which shows *-nx- ~ -nx^{w}-*. There is, moreover, a nagging similarity to Latin *sēbu* 'tallow, suet, fat, grease' (whence Spanish *sebo* and French *suif*), itself borrowed as Lapurdian *seba* (from the plural *sēba*). If borrowed early, this could come close to yielding this form, the main difficulty being that *-e-* rather than *-i-* would be expected. This source, or contamination from it, might explain the *-g-* in Lapurdian (including Baztán) *zigo*, which is not usually present in variant forms when *-h-* comes from **-n-*; note further that the following *o* might have encouraged a fluctuation between *-b-* and *-g-*, quite as in forms *aho ~ ago ~ abo* 'mouth' [11.].

[308.] Bq **zilar** 'silver'. Most agree that this word is a borrowing into Basque, probably ultimately from Akkadian *ṣarpu* 'refined silver'. Since the word can be projected back into Proto-Germanic, and also has Baltic and Slavic counterparts, it probably came to the Basque from northern Europe. After all, Gothic *silubr* resembles the Basque word more than does the Berber word cited. Tovar (1970:271-72) also prefers this connection, but sees borrowing as having gone in the opposite direction, out from the Iberian Peninsula, given the early attestation there of the exploitation of silver. The Roncalese and Bizkaian forms *zirar*, *zidar* also seem relevant to this question.

[312.] Bq **zu** 'you'. There is no prefixal form *zu-*, but rather just *z-*, as in *zara* 'you are', but there is a suffixal form *-zu*, as in *duzu* 'you have it'. Probably the *zu-* was taken by Bengston from the first part of words in which it is the stem, such as *zure* 'your'.

[313.] Bq **zuhain** 'hay, fodder', 'tree'. I question that this was **zunai*, which seems to be based on the presumption that we have present here a combining form *zun-* of *zur* 'wood' [315.]. We find a variant with *-g-*, *zugai*, in High Navarrese and Salacenco, and it is likely that this is a compound of the other combining form *zu-* + *gain* 'top, peak; upper surface; over, above' (cf. *uhain* [280.], *bekain* 'eyebrow' [my comment on 78.], and *mingain* 'tongue' [my comment on 218.]). We see *zu-* in examples like *zubi* 'bridge', where the second part is a reduced form of *bide* 'road, way' [67.], and *zuhaitz* 'tree', where the second part is probably *gaitz* 'difficult, etc.' Note additional meanings attested for *zuhain*: 'upper part of corn stalk' and 'second crop' (i.e., 'crop on top of another'). A nasal *-m-* is found only in *zumai* in the Baztán Valley, but nasalization could have spread from the word-final nasal, affecting an earlier glide **-w-*. This suggestion removes yet one more word as a separate entry.

5. Affixes

Affixes differ from stems in ways that make their separate treatment appropriate. They are usually unsubstantial in form, hence not susceptible to further analysis nor suspect of being borrowed, but lending themselves to accidental resemblances. In §4, Trask gives a thorough discussion of the not-quite-coterminous set of grammatical morphemes, so that a brief summary of the situation is all that is needed now.

Affixes ostensibly occur in 24 of the sets, but this must be adjusted by combining [145.] and [191.], and [227.] and [266.], noting the overlap of *-ik* (partitive suffix) between [160.] and [190.], adding *gu* 'we' [122.] with its implied affixes, removing secondary *-a* (article) [1.], discarding nonexistent **-ar* [23.] and **-tzi* [277.], and noting the occurrence of two morphemes, *-tzu* and *-tsu*, in [278.] (and two Caucasian morphemes in [160. 224.]). This leaves us with 20 sets involving 41 pairs. Of these pairs, 25 (61%) involve Caucasian languages, a greater proportion than for stems, implying heavy reliance on previously made comparisons to this group.

Pronouns occur in 5 sets involving 10 pairs [80. 122. 145. & 191. 227 & 266. 312.]. They appear as affixes in verb forms, both as prefixes and as medial and final suffixes. Nichols (1992:266) notes that in head-marking languages, such as Basque is, there tends to be a single pronominal root in independent pronouns, which may take verbal affixes. This is the case for Basque, except that there are two roots differentiated for number, taking the prefixes, singular *i* in *ni* 'I, we' and *hi* 'you', and plural *u* in *gu* 'we, us' and *zu* 'you'. But then in the medial verb forms for the singular pronouns, *-da-* and *-a-*, we can see a variant vowel *a*, so the singular pronouns are reconstructable as **di* ~ **da* 'I, me' and **gi* ~ **ga* 'you'. Some of the comparisons here are especially weak due to semantic differences: 'I' with Sumerian 'self' [227.], 'I' with Caucasian 'we' [266.], and 'you (pl.)' with Sumerian 'thou, thy' [312.].

For other verb morphology, only a causative prefix (with discrepancy of form) and a past tense suffix are compared [250. 225.].

From noun morphology, comparisons involve case suffixes [21. 155. 160. 190. 224. 250. 291.], plural number suffixes [189. 278.], and a suffix for indefiniteness [267.]. These constitute 10 sets with 24 pairs. Much as Trask notes, at least 9 pairs in 5 entries bring together markers of different categories [155. 160. 190. 224. 267.].

Derivational suffixes appear in 4 sets involving 8 pairs [202. 278. 290. 301.], most of which are semantically unlikely.

6. Taking Stock

In criticizing comparisons such as these of Basque to other languages, it seems useful to conceive of three categories of defects, with fuzzy boundaries between them. At one extreme we need to be sure that the word is genuinely a "native Basque word", as Trask (§2) uses the term. Thus one would exclude nonexistent words and morphemes resulting from modern coinage, misreading of sources, and misanalysis of longer forms, as well as probable borrowings from Latin or other sources, along with sound-symbolic words that respond to universal tendencies. At the other extreme would be characteristics of the actual comparisons made, which can be evaluated on both phonological and semantic grounds. In between is a large area, where a word might be genuine, but have an incorrectly applied or unrecognized analysis or reconstruction which weakens the particular comparison made. I will survey the reasons for unconvincing comparisons in this data, starting out from properties internal to Basque.

Most extreme is a set of 43 words or stems in 40 sets, entering into 74 pairs, that are nonexistent forms, to be set aside at first: **abe(t)s* 'voice' [3.], **atal* 'limb' [34.], **-koin* 'dog' [40.], **ba* 'son, child' [42.], **mik-usi* 'see, perceive' [55.], **-koki* 'forehead' [56.], **ber-* 'separate, distinguish' [60.], **beri* 'the same' [61.], **-hotz* 'heart' [70.], **-rika* 'lung' [72.], **-zi* 'life, alive' [74.], **-zka(r)* 'back' [75.], **bortz* 'hand' [76.], **-p(h)uru* 'eyebrow, eyelash, eyelid' [78.], **-kume* 'woman' [88.], **-zte* 'woman, wife' [89.], **eme* 'sweet' [90.], **-bar(r)-* 'palm' [97.], **gal-* 'side, armpit' [106.], **gara* 'skull', **-khotx(e)* 'bone' [109.], **gul-* ~ **gil-* 'egg (?)' [114.], **gor* 'meat' [119.], **-di* 'big' [129.], **-haur* 'self' [136.], **-gal* 'wing' [140.], **ika* 'one' [161.], **ilu* 'move' [163.], **-tzig-* 'frost' [166.], **kala* 'castle (?)' [194.], **lor-* 'track, etc.' [208.], **-kuts-* 'sleeve' [211.], **maño* 'masculine' [214.], **zor* 'two', **-tzi* 'ten' [277.], **-khab-* 'fist' [281.], **-kharai* (and variants) 'wrist' [282.], **-khondo-* 'elbow' [283.], **-zki* 'anus' [288.], **zaki* 'bone' [295.], **zaro* 'night' [299.], **zartzu*, **ziri* 'sharp' [311.].

The borrowings into Basque from the Latin-Romance continuum are of course very numerous, and need to be distinguished from native words (these have been studied most recently by Mujika 1982). In this context, one is always reminded of A. Griera (1960), who apparently assumed that Basque originated from Latin, offering many etymologies without undue concern for phonological or semantic constraints. I report a baker's dozen for words in this paper's sets (these are in the same relative order; Griera gives Southern forms lacking *h*; I have added translations of Latin forms): *ao* 'mouth' < BUCCA (with metathesis), *esku* 'hand' < DISCU 'quoit', *etxe* 'house' < TECTU 'roof; dwelling', *gari* 'wheat' < GRANU 'grain, seed', *gorri* 'red, etc.' < RUBEU, *andi* ~ *aundi* 'big' < GRANDIS, *otz* 'cold' < FRIGIDUS, *izen* 'name' < DICENTE 'saying', *sagu* 'mouse' < CAECU 'blind', *sudur* 'nose' < IPSU ODORE 'with the smell itself', *uda* 'summer' < VERE 'in the spring', *ur* 'water, juice' < AQUA, *zaar* 'old' < SENARIU.

For those with a more critical approach to the question, there are often difficulties in reliably identifying Basque words from this source, so one is again reminded of the continuum of degrees of probability that obtain in etymological investigations. One can be most certain of this source if the word is attested in early Latin documents or has cognates occurring at some geographical remove from the Basque-Iberian area. But there clearly have been importations into peninsular and Pyrenean Romance languages from Basque or Iberian substratum. The most notorious are certainly the words related to Basque *ezker* 'left', which include Castilian *izquierdo*, Portuguese *esquerdo*, and similar forms in Catalan, Gascon, and Occitan; several etymologies within Basque have been suggested, most of them starting with *esku* 'hand', such as **esku oker* 'crooked hand' and **esku erdi* 'half hand (i.e. half of the pair)' (Hubschmid 1953-57, Corominas 1954[2]:1014-17, Tovar 1959:26-31). Among the Castilian words mentioned in Trask's discussions, the following strike me as most likely to be of similar indigenous origin: *cuscorro* 'hard bread crust' [239.], *binza* 'thin skin on the body of an animal' [245.], and *zarza* 'bramble' [257.], less certain is *mogote* 'pointed stack of corn, etc.' [55.]. There are also probable borrowings from Celtic, Gaulish, or Celtiberian, into Basque as well as into surrounding Romance languages, one of which may be Castilian *gancho* 'hook' [193.].

I have tried with some uncertainty to distinguish the earlier borrowings from Latin from later ones from Romance. From Latin, there are 21 words entering into 33 pairs: *aberats* 'rich' [2.], *aditu* 'hear, listen, understand' [6.], *agure* 'old man' [8.], *beko* 'face' [55.], *biga* 'two-year-old heifer' [68.], *busti* 'wet, moist' [79.], *gela* 'room, chamber' [111.], *golko* 'space between one's chest and one's clothes', 'bay, gulf' [118.], *gurpil* 'wheel', *inguru* 'vicinity, around, near' [125.], *horma* 'ice', 'wall' [151.], *tiña* 'ringworm' [170.], *kaiku* 'wooden cup or bowl' [192.], *mauka* 'sleeve' [211.], *matel* 'cheek' [217.], *pataxa* 'bottle' [243.], *pipil* 'bud' [246.], *ondo* 'bottom, below, next to' [283.], *xahu* 'clean' [289.], *zamar* 'sheepskin, sheepskin jacket' [297.], *zuku* 'soup, broth' [314.]. Two of these enter into misanalyzed words of a previous paragraph [211. 283.]. An additional Latin borrowing occurs, *eme* 'female', but either is not the part of the word that is compared [88. 89.] or was compared with an erroneous meaning [90.].

From Romance, there are 25 words in 21 sets, compared in 42 pairs: *akain* 'tick' [14.], *apo* 'toad' [22.], *arroda* 'wheel' [30.], *barakurkuilo* 'snail' [44.], *beginini(ko)* 'pupil' [52.], *galtzar* 'side of the body' [106.], *ganga* 'palate', *gangar* 'uvula' [107.], *krako* 'hook' [193.], *kaloi(a)* 'hut, cabin' [194.], *kankano* 'large, clumsy person' [195.], *kasko* 'skull, crown' [196.], *kokot* 'nape' [198.], *kurlo* 'crane' [201.], *mama* 'breast', 'mother' [213.], *mulo* 'pile, stack (of hay or cornshocks)' [221.], *musu* 'kiss, face', *musin* 'snout, muzzle' [222.], *paru* 'stick of wood, pole' [242.], *potorro* 'vulva' [247.], *senton* 'old man', *sentana* 'old woman', *santan* 'ever' [259.], *toska* 'kaolin', Z. *toxka* 'clod' [268.], *zango* 'leg' [298.].

There are also four Spanish words themselves entered in sets. Two of these are in addition to an attested Basque word, no doubt implying borrowing from it: *zorro* 'fox' beside *azeri* ~ *axeri* [39.], and *cachorro* 'puppy' beside *txakur*, *zakur* 'dog' [296.]. These connections are questionable, but I would not completely dismiss the latter. The two other Spanish words have no attested Basque counterpart: *marrano* 'pig' [216.], which seems to have come from Arabic, and *becerro* 'bullock' [50.], which may indeed be an indigenous word.

In his discussions for eight of the sets [26. 33. 40. 46. 62. 77. 132. 162.], Trask reports conjectured borrowings from branches of Indo-European other than Italic, especially Celtic and Germanic, or just unidentified early Indo-European. This seems most attractive in the case of animals, *hartz* 'bear', found by some also in *asto* 'donkey' and *azkoin* 'badger', and *bargo* 'young pig', and an animal product, *ile* 'hair, wool', but the evidence is not strong enough to preclude any of these words from being native Basque.

There are also sets embodying suggestions of borrowing into Basque from Afroasiatic, five from Berber [67. 178. 181. 239. 308.], and three from Ancient Egyptian as represented by Coptic [27. 63. 300.], and also from Kartvelian (South Caucasian), as represented by Georgian [193.] or Svan [239.], plus one that might be from either Berber or Georgian/Laz [285.]. Two other sets suggest borrowing from Dene-Caucasian into Indo-European: 'wine' into Albanian and Armenian [25.] and 'hunting dog' into Sardinian and Greek dialect [296.]. These are all entirely unconvincing, but can be set aside as not concerning our central genetic hypotheses. Along with the Spanish words already mentioned, these will remove 15 sets and 20 pairs from further consideration.

In his comments on many of them, Trask rightly points out that it is inappropriate to include words embodying sound symbolism, in the broad sense, in comparisons suggesting distant relationship, since, given the limited arbitrariness of the connection between sound and meaning that defines them, they are constantly being re-formed. In these data, there are two straightforwardly onomatopoetic words: *t(h)u ~ ttu* 'spit' [171. 269.] and *uzki* 'anus, buttocks' (based on *utz* 'breaking of wind', variant of *putz* 'puff of breath, fart' [288.]; additionally *txori* 'bird' [275.] might suggest chirping, but this is weaker. Some nursery words, arising from the imitation of infant babbling, also occur: *mama* 'liquid' [212.], *mama* 'breast', 'mother' [213.], *ama* 'mother' [18.]. Suspect because of various combinations of reduplicative pattern and variant forms are: *kankano ~ kankanu ~ kankan* 'large, clumsy person, etc.' [195.], *maguri* 'strawberry' [210.], *mara-mara* 'smoothly' [215.], *pinpirin* 'butterfly' [244.], and *tximeleta* 'butterfly' [273.] (the last two probably also having Romance connections); one might include here also *zizari ~ zizare ~ xixari ~ txitxari* 'worm' [309.]. Sound symbolism of [i] suggesting smallness is certainly present in *txiki ~ txipi ~ ttipi* 'small' [272.] and the *-nini-* of *beginini(ko)* 'pupil (of the eye)' [52.]. Related shapes suggesting roundness (like English *ball*) are seen in the *-bil* in *barrabil* 'testicle' [47.], *gurpil* 'wheel' [125.], *ukabil* 'fist' [281.]; the *-bor* of *zilbor* 'navel' [307.]; the *gil- ~ gul-* of *giltzurrin ~ gultzurrin* 'kidney' [114.]; such probably also obtains in the *gur-* of *gurpil* 'wheel' and *inguru* 'vicinity' [125.], which is probably borrowed; lastly, such shapes tend to occur in words for 'boil (v.)', because of the round bubbles, as Trask points out regarding the Burushaski word paired with Basque *bero* 'hot' [62.]. Four of these were listed above as borrowings; we may now exclude another 9 sets with 19 pairs.

With these factors, we have now removed from consideration many of the non-native Basque words, although certainly many more occur in the remaining sets. We can now list some 28 words entering into 51 pairs, for which a better morphemic analysis, etymology, or reconstruction seems to make the comparison unlikely: *ahal* 'ability' [9.], *ahizpa* 'sister (of a woman)' [10.], *ahuñe* 'kid' [12.], *ahur* 'palm (of the hand)' [13.], *alderatu* 'separate, remove' [16.], *arraultza* 'egg' [29.], *axuri* 'young lamb' [37.], *barakurkuilo* 'snail' [44.], *barrabil* 'testicle' [47.], *behazun* 'bile' [54.], *bihar* 'tomorrow' [69.], *ehiza* 'hunting' [85.], *haitz* 'rock, crag' [127.], *haragi* 'meat, flesh' [130.], *hazi* 'grow, grow up' [139.], *ihintz* 'dew' [159.], *intzaur* 'walnut' [165.], *ithoitz* 'roof, gutter' [171.], *jin* 'come' [186.], *ohoin* 'thief' [232.], *txahal* 'calf' [270.], *txar* 'bad' [271.], *txingurri* 'ant' [274.], *txori* 'bird' [275.], *txorru* 'root (of a hair)' [276.], *uhain* 'wave (n.)' [280.], *uxuri* 'urine' [287.], *zuhain* 'hay, fodder; tree' [313.].

Having made these exclusions, we have reduced our corpus to some 170 sets involving 365 pairs. We now turn to characteristics of the comparisons, which I will group into phonology, semantics, and topology of sets. Here I forego focusing on acceptance or exclusion of individual items and consider factors which affect the probability of these comparisons taken as a whole.

In assembling these sets, it is clear that the compilers did not apply any theory of sound correspondences, in the sense of excluding comparisons because they did not fit. Rather, general criteria of phonetic similarity have been applied. Yet one should not be excessively strict about sound correspondences when it comes to very distant relationships, since the loss of evidence of conditioning factors will make what were originally regular conditioned sound changes come to seem irregular.

In §4, Trask tabulates the Caucasian and Burushaski correspondences for Basque *z-* and *s-*. In fairness, it should be pointed out that the situation is not quite as diversified as he suggests. For *s-*, **sagar** and **soin** show the same correspondence, and since absence of a cognate cannot be taken as counterevidence, **sahats** can be grouped with any of four others. For *z-*, **zain** and **zorri** are the same, and should be grouped with **zer**; **zakar** can go with either **ziho** or **zikiro**; **zuzen** should probably be regarded as showing Burushaski *ch-*, so it can go with either **zahar** or **zikiro**; **zorrotz** [311.] shows another case of the same correspondence as **zakar**. But for potential additional corresondences, see [301. 307.] and my comments to [271. 275.]. It is, however, unrealistic to expect three-way

regular correspondences. Just in comparing pairings with Basque, we find for Burushaski 4 or 5 with *s-* and 6 or 7 with *z-*; and for Caucasian, 11 with *s-* and 10 or more with *z-*.

One other desideratum to consider is that comparisons should be based on matching several phonemes in a word, when possible, probably at least *CC* or *VCV*. It seems that in gathering these sets, a *CVC* matching was in most cases attempted. Of course, this is unavailable in the case of Basque *CV* stems such as *bi* [66.], *lo* [207.], *su* [263.], *no-* [229.], and *ze-* [302,], and *VC* stems such as *ar* [24.]; for them, inevitably the chances of accidental matchings are greater. In the languages compared, monosyllabic stems are especially frequent in Northwest Caucasian and Sumerian. Due to overlap in matched forms, it is fairly often the case that no more than this amount of longer stems is the actual basis for connecting a pair. In a few pairs, there is really only one shared phoneme, such as the consonants in Bq. *itz* : Abkh. *cʷa* [149.] and Bq. *no-* : PY. **ʔan-* [229.]. It is also noteworthy that vowel qualities, rather than the mere presence of a vowel, are at best sporadically relevant to the comparisons at hand.

Turning now to meanings, a glaring weakness of these comparisons is what I call semantic discontinuity: the bringing together in comparisons of forms whose meanings differ to the extent that most languages would have separate words for them. Trask certainly comments on the semantic implausibility of many sets, and offers appropriate corrections of the meanings of many Basque words.

My reservation here is not to be misunderstood as a claim that change between such meanings is impossible. Thus, although I find semantic discontinuity in these pairs: Bq. 'dream' : PY. 'sleep' [19.], Bq. 'sea' : Cauc., PY. 'salt' [172.], and Bq. 'moon' (actually a compound with 'light') : Cauc. 'sun' [26.], I am aware of Spanish *sueño* 'sleep, dream', Makah *tupʼaƛ* 'salt, ocean', and Washo *díbe* 'sun, moon', that combine such meanings in a single word.

The problem is that being semantically lax in this way opens up too many possibilities for chance juxtapositions. Thus, a set like [15.], which pairs Basque 'daughter' with Hurrian 'sister', has doubled the possibility of a word's being matched; one like [35.], which allows for a relationship among Basque 'old woman', Caucasian 'female', and Burushaski 'paternal aunt', has tripled the possibility (actually more, since other meanings such as 'maternal aunt' and 'grandmother', are implied as relatable), and so forth.

Such discontinuities occur in 146 pairs in 84 of these sets, i.e., about half of the sets and two-fifths of the pairs. It is therefore pointless to list them all, but here are 20 that are clear enough cases: Bq. 'age' : Bur. 'year' [5.], Bq. 'dry' : Bur. 'very old and wrinkled' [7.], Bq. 'mouth' : Sum. 'eat' [11.], Bq. 'spleen' : Sum. 'liver' [45.], Bq. 'middle, half' : Bur. 'two' [92.], Bq. 'finger' : Cauc. 'cubit', 'armful, bosom' [93.], Bq. 'come' : Cauc. 'run' [99.], Bq. 'hard' : Cauc., Bur. 'stone' [117.], Bq. 'bear' : PY. 'badger' [132.], Bq. 'breath' : Cauc. 'odor' [133.], Bq. 'dead', 'die', 'kill' : Bur. 'army' [146.], Bq. 'twenty' : PY. 'ten' [150.], Bq. 'ox' : Cauc. 'ram' [157.], Bq. 'star' : Sum. 'new moon' [176.], Bq. 'mane' : Bur. 'felt' [200.], Bq. 'friend, companion' : Cauc. 'slave' [203.], Bq. 'sleep' : Sum. 'lie' [207.], Bq. 'tear (n.)' : Bur. 'boil (n.)' [228.], Bq. 'shoulder, body, torso' : Cauc. 'game animal' [262.], Bq. 'summer' : Sum. 'sun, day' [279.]. Some sets contain additional words that seem intended as semantic bridges; for example, in [93.] there is also Bur. 'hand', in [146.] also Sum. 'destroy', in [176.] also Bur. 'morning star (Venus)', and in [228.] also Cauc. 'tear, pus'. These are nevertheless artifacts of the comparer that should not be taken to justify the discontinuities.

Equally serious limitations are seen in the patterns of representation of languages or families in the comparative sets. Although Trask (§1) mentions only "Macro-Caucasian", subsuming Basque, North Caucasian, and Burushaski, representatives of five other Dene-Caucasian groups turn up in the sets: Ancient Near Eastern Sumerian, Hurrian, Urartean, and Hattic, and Siberian Yeniseian. Thus, we are presented with hypotheses of the relationship of Basque to seven other groups, those other than North Caucasian and perhaps Yeniseian being single languages. Three of the Near Eastern groups occur in few sets, so a relationship to them cannot be considered demonstrated: Hattic in 4, Urartean in 5, and Hurrian in 6. About the same number of sets include Yeniseian and Sumerian, 44 and 48 respectively. Burushaski appears in 90, and Caucasian in by far the most, 168.

Very significant is the fact that 76 sets, 45% of the total, involve the pairing of Basque with only one other group, as follows: Caucasian 51, Burushaski 11, Sumerian 8, Yeniseian 4 (one also including Proto-World), Hurrian 1, and Proto-World 1. One acknowledges that some comparisons have been presented elsewhere among these other languages, but still we are seeing a version of the approach of declaring a group of languages to be a unit and treating comparisons to any one of them as though to all.

Trask (§4 end) appropriately raises the question of the reliability and handling of the data of the other languages. Some observations can readily be made for the featured North Caucasian concerning the pedigree of forms compared, i.e., the likelihood that they were already present in the proto-language of the family. North Caucasian is deeply divided into Northwest (NW) and Northeast (NE) branches, the latter in turn between a small North Central (NC) and the still highly diversified remaining NE groups. Note that a source such as Geiger, et al. (1959:9), intending to be a reference of agreed-upon findings, lists three separate Caucasian families, NW, NE, and Southern (Kartvelian). Thus one would hope to have a compared form confirmed by cognates between NW and NE. But in 36, or 71%, of the sets, the Caucasian form comes from only one of the branches: 16 from NW, 2 from NC, and 18 from NE. Moreover, 10 forms are taken from single languages: 6 from NW Abkhaz, 1 from NC Batsbi, 1 each from NE Archi, Budukh, and Tsakhur.

Based on all of these considerations, it should be eminently clear that the set of comparisons under consideration does not confirm the hypothesis of a genetic relationship of Basque to any of the seven other families or language isolates with which it was paired.

Additional References

Bouda, Karl. 1952b. "Awarisch *x*, tscherkessisch *L* und der baskische Rhotazismus." *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft* 4:252-8.

Bouda, K., and D. Baumgartl. 1955. "Nombres vascos de las plantas." *Acta Salmanticensia Filosofía y Letras* 7(3):1-99.

Braun, Jan. 1981. "Esukaro - Caucasica." *Euskalarien Nazioarteko Jardunaldiak*, pp. 212-19. (Iker - 1.) Bilbao: Euskaltzaindia.

Castaños Garay, Florentino. 1979. *La formación del plural en la Lengua Vasca.* Bilbao: Diputación Foral del Señorío de Vizcaya.

Corominas, J. 1954. *Diccionario crítico etimológico de la lengua castellana.* Berne: Editorial Francke.

Corominas, Joan. 1973. *Breve diccionario etimológico de la lengua castellana.* 3rd ed. Madrid: Editorial Gredos.

Gavel, H. 1921. "Éléments de Phonétique Basque." *RIEV* 12:1-543.

Geiger, Bernhard, Tibor Halasi-Kun, Aert H. Kuipers, and Karl H. Menges. 1959. *Peoples and Languages of the Caucasus.* (Janua Linguarum, 6.) The Hague: Mouton & Co.

Gensler, Orin David. 1993. *A Typological Evaluation of Celtic/Hamito-Semitic Syntactic Parallels.* Ph.D. Dissertation, University of California, Berkeley.

Griera, A. 1960. *Vocabulario Vasco (ensayo de una interpretación de la lengua vasca).* Abadía de San Cugat del Vallés: Instituto Internacional de Cultura Románica.

Gursky, Karl-Heinz. 1974. "Der Hoka-Sprachstamm. Eine Bestandsaufnahme des lexikalischen Beweismaterials." *Orbis* 23:170-215.

Hoz, Javier de. 1981. "El euskera y las lenguas vecinas antes de la romanización." *Euskal Linguistika eta Literatura: Bide Berriak*, pp. 27-56. (Euskal Herria, 4.) Bilbao: Universidad de Deusto.

Hubschmid, Johannes. 1953-57. "Vasco *ezker* 'izquierdo.'" *Eusko-Jakintza* 7:97-100.

Jacobsen, William H., Jr. 1979. "Hokan Inter-Branch Comparisons." *The Languages of Native America: Historical and Comparative Assessment*, ed. by Lyle Campbell and Marianne Mithun, pp. 545-91.

Kuipers, Aert H. 1960. *Phoneme and Morpheme in Kabardian (Eastern Adyghe).* (Janua Linguarum, 8.) The Hague: Mouton & Co.

Lafon, René. 1947. "L'état actuel du problème des origines de la langue basque." *Eusko-Jakintza* 1:35-47, 151-63, 505-24.

Martinet, André. 1950. "De la sonorisation des occlusives initiales en basque." *Word* 6:224-33.

Martinet, André. 1955. "La reconstruction structurale: Les occlusives du basque." *Économie des changements phonétiques: Traité de phonologie diachronique*, pp. 370-88. Berne: Éditions A. Francke.

Martinet, André. 1981. "La phonologie synchronique et diachronique du basque." *Euskalarien Nazioarteko Jardunaldiak*, pp. 59-74. (Iker - 1.) Bilbao: Euskaltzaindia.

Mujika Urdangarin, Luis Maria. 1982. *Latina eta Erromanikoaren Eragina Euskaran (Euskal lexikoaren azterketa bideetan).* San Sebastián: Sendoa.

Nichols, Johanna. 1992. *Linguistic Diversity in Space and Time*. Chicago: University of Chicago Press.

Tovar, Antonio. 1959, *El Euskera y sus parientes*. (Biblioteca Vasca, 2.) Madrid: Ediciones Minotauro.

Tovar, Antonio. 1970. "The Basque Language and the Indo-European Spread to the West." *Indo-European and Indo-Europeans*, ed. by George Cardona, Henry M. Hoenigswald and Alfred Senn, pp. 267-78. Philadelphia: University of Pennsylvania Press.

Uhlenbeck, C. C. 1942. *De oudere lagen van den Baskischen woordenschat*. Mededeelingen der Koninklijke Nederlandsche Akademie van Wetenschappen, afdeeling Letterkunde, Nieuwes Reeks, Deel 5, No. 7. Amsterdam. French translation (1947), "Les Couches Anciennes du Vocabulaire Basque." *Eusko-Jakintza* 1:543-81.

Uhlenbeck, C. C. 1949. "Los nombres vascos de miembros de cuerpo que comienzan con *b-*". *Eusko-Jakintza* 3:105-11.

Urreiztieta-Rivera, Irel. 1980. *Basque and Caucasian: A Survey of the Methods Used in Establishing Ancient Genetic Affiliations*. Ph.D. Dissertation, The University of Arizona.

Voegelin, C. F. and F. M. 1959. "Guide for Transcribing Unwritten Languages in Field Work." *Anthropological Linguistics* 1(6):1-28.

Waterhouse, Viola Grace. 1976. "Another Look at Chontal and Hokan." *Hokan Studies: Papers from the First Conference on Hokan Languages Held in San Diego, California, April 23-25, 1970*, ed. by Margaret Langdon and Shirley Silver, pp. 325-43. The Hague and Paris: Mouton.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (November 1995)

COMMENTS STIMULATED BY L. TRASK'S CRITIQUE

Edwin G. Pulleyblank
University of British Columbia
e-mail: edwin@unixg.ubc.ca

Professor Trask's critique of the proposed connection between Basque and Dene-Caucasian seems to me to be an admirable example of the kind of searching scrutiny that needs to be applied to other facets of the Dene-Caucasian hypothesis and to other long-range language comparisons that have become fashionable in recent years and that have been promoted in the pages of *Mother Tongue*. In saying this, I want to emphasize that I am not against the idea of such long-range connections in principle. It is a question of what constitutes reliable evidence and how one goes about testing hypotheses.

Merritt Ruhlen (1994:12-13) claims that Sir William Jones's proposal in 1786 that Sanskrit, Greek, and Latin had "sprung from a common source" and that Gothic and Celtic might also have had the same origin constitutes an "evolutionary explanation for linguistic diversity ... discovered some time before Darwin's parallel discovery of evolution by natural selection in biology." This seems to me fallacious on a number of counts. In the first place, it assumes that there is a real parallel between the way in which languages change over time and biological evolution. In the second place, it confuses the idea of the evolution of species over time, which long preceded Charles Darwin, with his hypothesis, supported by meticulous observations and close reasoning, that the mechanism of evolution was natural selection rather than some other mechanism such as the inheritance of acquired characteristics. In the century and more since the publication of *The Origin of Species*, this hypothesis has been subjected to rigorous testing on all sorts of evidence and has been deepened and expanded by other discoveries, including Mendelian inheritance and the biochemical structure of genes, in spite of which many aspects remain subject to vigorous debate.

In spite of the common metaphor which speaks of languages as having "ancestors" and being "genetically related", there is no close analogy between biological evolution and the ways in which languages are transmitted from one generation to the next and gradually change in the process. Languages do not have ancestors. They have speakers. The speakers have ancestors but speakers do not inherit their language through their genes. They acquire their language by growing up in a community where the language is spoken. There is indeed something genetic involved as we can tell from the fact that only humans can acquire a human language. What this innate capacity for language is, how it is transmitted in our genes, and how it has developed in the course evolution are fascinating and difficult questions, but they are only very indirectly related to what is ordinarily meant when one speaks of genetic relationships between languages.

That languages change over time was known long before Sir William Jones gave his famous lecture and all sorts of speculative theories abounded, for instance, about which language was closest to the common language of mankind before the confusion of tongues at the Tower of Babel. His real contribution was in focusing attention on the lexical and morphological similarities between the classical languages of Europe and India which, it turned out when they were systematically studied, could be accounted for by the further hypothesis that the sound systems of the languages in question had changed in systematic ways independently of the words which they encoded. This doubly articulated nature of language is why language change is different from other cultural changes that take place over time. This is why we can say with confidence that words as different as *cow* and *beef* in English go back by different routes to a common origin thousands of years ago.

When first enunciated, the Neogrammarian principle of the regularity of sound change was an empirical discovery without a solid theoretical foundation. A first step in providing such a foundation was the Saussurean concept of the arbitrariness of the linguistic sign and the idea that at any given synchronic stage the sound system of a language was made up of a set of mutually contrastive "phonemes". Sound change could then be summed up in Bloomfield's pithy sentence: "phonemes change" (1935:351). Unfortunately, emphasis on the phoneme as a kind of atomic unit defined not in terms of its phonetic content but in terms of the network of contrasts into which it entered in the sound system of the language had the effect of making it difficult to understand in a principled way how one

synchronic system could change into another system with a different number and arrangement of phonemic "atoms" (Allen 1953). Since the fifties, the theory of universal distinctive features that emerged in the Prague school, especially the work of Roman Jakobson, has become a basic premise of the generative school. It breaks down, so to speak, the phonemic "atom" into more basic units and makes it possible to treat diachronic phonology as a part of the same discipline as synchronic phonology. Phonetic change consists in the rearrangement of phonetic features that are already present and universally available rather than the substitution of a new set of atomic phonemes. Unfortunately, I do not see much evidence of interest in these recent developments in the work of the new "long-rangers" or in the compartmentalized studies of historical linguistics in Chinese, Sino-Tibetan, and Indo-European with which I am more or less familiar.

From the beginning, the achievements of Jones and his successors in establishing the Indo-European language family naturally led to attempts to apply the same methodology to other sets of apparently related languages such as Semitic, Uralic, Altaic, and Sino-Tibetan and to make connections between Indo-European and other languages and families of languages. Some of the families postulated in the nineteenth century are now as well established as Indo-European. Others remain in dispute. These include, for example, Altaic, which some wish to break up into its three most widely accepted components, Turkic, Mongolian, and Tungusic, and others wish to extend to include Korean and Japanese. Sino-Tibetan was once assumed to include Tai and Miao-Yao, and this is still the prevailing view among Chinese linguists, but most western linguists have accepted Paul Benedict's view (1942) that the many lexical contacts between these languages and Chinese are the result of comparatively recent borrowing, not inheritance from a common source. Even the connection between Chinese and Tibeto-Burman, which is accepted by Benedict and seems solid to me in spite of the limited success that has so far been achieved in establishing a large body of cognates linked by regular sound correspondences, is disputed by Roy Andrew Miller (1988) (the most ardent advocate of including Japanese in Altaic!) and Lauren Sagart (1994) (who wants to link Chinese instead with Austronesian).

In a situation like this, there is obviously a need to find some common ground as to what constitutes proof. It is claimed by the new long-rangers that they have a new methodology that renders older objections to the kind of claims they make obsolete. Just what this new methodology is is not clear to me. In the context of the present discussion, the best I can offer is an examination of the claims of Sergei Starostin to have established a connection between Sino-Tibetan and Northern Caucasian, with Yeniseian as an adjunct or intermediary, constituting one of the main components of Dene-Caucasian. Let me start with his remarks on "Methodology of Long-range Comparison" (1988). He first offers two arguments to justify the existence of such studies: (1) the need for some way of classifying the many different language families that are currently recognized, (2) the fact that comparative linguistics is one of the few ways of investigating the preliterate history of mankind. This is, of course, about motive rather than method. The real questions are whether and how these desirable ends can be achieved. The first point he makes about method as such is that, in contrast to traditional comparative linguistics which "relies basically on written and spoken languages ... the basic material for long-range comparison is reconstructions." To illustrate the need for reconstruction, he lists modern forms in Chinese and Burmese for the numerals '2', '5', and '8', which seem to have nothing in common, whereas reconstructed forms bring them together. (He does not list any of the many other attested Sino-Tibetan forms that make it probable that the words in question are related even without a reconstruction and which provide the basis for the reconstructed forms he cites.) He then cites more problematic cases of alleged correspondences between reconstructed forms in Indo-European and Altaic and between reconstructed forms in Chinese and Caucasian. He claims that reconstructed Proto-Altaic **k'üjla* 'ear, to hear' is closer phonetically to Proto-Indo-European **ḱleu-* 'to hear' than Turkish *kulak* 'ear', Korean *kwi* 'ear', Evenki *ūl-ta* 'to be heard, resound' are to Russian *slyžať* 'hear', Old Indian *śru-* and English *loud.* True, but what about Greek κλύω 'hear' which might suggest a relationship to Turkish *kulak* all by itself? With or without the reconstruction, I don't know whether the Altaic and Indo-European forms really go back to a common source. What I, as a traditionally minded historical linguist, would like to see is *patterns* of sound correspondences drawn from semantically unrelated sets of words that would justify the assumption that it is right to ignore the *discrepancies* between the Altaic and Indo-European forms — the aspiration of the Proto-Altaic initial consonant (if that is what is

meant by **ḱ*), the **-j-* inserted after the nuclear vowel and the difference in the order of the liquid and the back-rounded vowel/glide.

Turning to the question of Chinese and Caucasian, there is an obvious contrast with the Indo-European in that, while there is a large body of Indo-European reconstructed forms that are well understood and generally accepted (though there are still many problems about such things as the so-called "laryngeals"), the same is certainly not true for Sino-Tibetan or even Old Chinese. (I say nothing about North Caucasian, since I have no idea how well Starostin's reconstruction is accepted by others working in that area.) This can be illustrated by the third example he cites in the above-mentioned article to illustrate his point about the use of reconstructed forms. He claims that Chinese *yi* 'breast' and Burmese *raŋ* breast' can be related to Chechen *dog* 'heart', Agul *jirḳʷ* 'heart', Circassian *gʷe* 'heart, breast' through Proto-North-Caucasian **jerḳʷi* and Proto-Sino-Tibetan **ʔrək/*ʔrəŋ*. Even if we accept the reconstructed forms as well established, the phonetic correspondence boils down to (possibly) sharing two consonantal phonemes, **r* (attested in Agul and Burmese) and some kind of back consonant. The **r* in the Chinese is, however, quite uncertain — in my opinion quite spurious. Starostin himself implies this uncertainty elsewhere when he reconstructs the word as **ʔ(r)ək* (Starostin and Ruhlen 1994:73). In his reconstruction of the rhymes of the Book of Odes, the earliest internal source that he uses for reconstructing Old Chinese, he reconstructs the homophone *yi* 'large number, hundred thousand' simply as **ʔək* (1989:552). The insertion of **-r-* seems to be justified only by (a) the wish to compare the Chinese and Burmese words, and (b) the theoretical assumption (which would take too much space to discuss here) that, if present, a medial **r* would have been lost without trace. As I have recently argued, there is, in fact, good reason to think that initial glottal stop in Old Chinese was not a true consonant but simply an obligatory onset for vowel initial words and that **ʔr-* clusters were impossible (Pulleyblank 1995). If we eliminate **r* from the Chinese form, Burmese *raŋ* is left as the sole representative of the supposed Tibeto-Burman proto-form (which does not appear in Benedict 1972). The claimed "plausibility" of the comparison with Proto-North-Caucasian is not apparent to me.

Many of Starostin's other proposed Sino-Caucasian cognates that involve Old Chinese reconstructed forms seem to me to be open to the same kind of objections. I have discussed those cases in another recent paper (Pulleyblank forthcoming b). I will not repeat these here but will add a few examples taken from Starostin and Ruhlen (1994):

AWL: Proto-Yenisei **duʔt*, Ket *duʔt*, Yug *duʔt*, Kott *tūt*. Cf. Proto-Abkhaz-Adyg **ɗʷəɗʷə*, Proto-Abkhaz-Tapant **ɗᶻaɗᶻə*, Old Chinese *tuj*. [No other Sino-Tibetan form cited.]

For reasons set out in Pulleyblank (1994), I would reconstruct this word as **kʷjəl*. No doubt Starostin would disagree, but he needs to account for the rest of the extensive phonetic series to which the graph in question belongs. In the absence of any other Sino-Tibetan form, **tuj* 'awl' does not seem to be a strong candidate for comparison with other languages.

BEAR (n.): Proto-Yenisei **čajaŋ*, Kott *šajaŋ*, Pumpokol *xanxi*. Cf. Proto-North-Caucasian **cʰʷäʔnʌ*, Proto-Tibeto-Burman **s-wam*, Old Chinese **wəm*, Burushaski *yã*.

Benedict (1972:116) reconstructs the Tibeto-Burman form as **d-wam*, with **d ~ *t* as a (syllabic) prefix. He does also suggest that Kanauri and Thebor *hom* may be for **s-wam* instead of **d-wam* (with **s-* as an animal prefix) and also quotes Mikir *si-tum* < **twam* as combining the two prefixes. It seems clear that the basic root is **wəm* or **wam*. The only correspondence to this in Yenisei forms is the final (velar!) nasal. Proto-North Caucasian at least has a labialized initial but still seems to require an affricate that can only be found in Sino-Tibetan by including **s-* as an integral part of the root. This does not seem to me to be a good candidate for a cognate shared by Sino-Tibetan, Yenisei, and Caucasian.

BILE / BITTER: Proto-Yenisei **qʌqʌr*, Ket *qʌ̄ɬ* 'bile', *qōlɨŋ*, Yug *xʌxul* 'bile', *xʌxɨɬaŋ* 'bitter', Kott *ogar* 'bile', ?Pumpokol *leō-xoxar* 'bladder' (? < 'gall bladder'). Cf. Proto-North-Caucasian **q'eq'ʌ ~ *-eq'ʌ-* 'bitter', Proto-Tibeto-Burman **ka* 'bitter', Old Chinese **qāʔ*, ...

There is no problem here about the equivalence of the Chinese and Tibeto-Burman forms (on which see Benedict 1972:18, with other references throughout the text) and the semantic equivalences seem very good. What is not clear to me is the basis for the reconstruction of the Old Chinese with a uvular stop, especially since a plain velar is assumed for the Tibeto-Burman form. Elsewhere Starostin reconstructs the Chinese word as **khāʔ* (1989:561). The correspondence of Chinese aspirated **kʰ* to Tibeto-Burman **k* as reconstructed by Benedict, is regular. The word in question if *kha* in Tibetan. Though the phonetic correspondence linking the Sino-Tibetan, Yenisei, and Caucasian forms reduces itself to having some kind of back consonant as the initial, the semantic agreement seems impeccable and, *if* the languages in question are indeed genetically related, this may be a good candidate as a part of the common vocabulary.

WHO_2: Proto-Yenisei **ʔan-*, Ket *anej* ~ *ana*, Yug *anet*. Cf. Proto-(North)-Caucasian **nʌ* (interrogative pronoun base), Sino-Tibetan: Old Chinese **nāj* 'how, what', Tibetan *na* 'when', Khamti *an-nan* 'what', Nahali *nāni, nān* 'what'.

The Old Chinese form is not one of the usual interrogative pronouns but a rare particle that may be a contraction of *rú hé* < **nà gál* 'like what' (= **na gāj* in Starostin's reconstruction) (Pulleyblank 1988:349). The initial **n-* is found in an extensive word family that includes other words meaning 'like, thus, so, etc.' but no interrogatives except secondarily, as in this case. Tibetan *na* 'when' has equally dubious credentials. According to Jäschke's *Dictionary*, it is a locative postposition meaning 'in, at, with, to' which, with verbs, is equivalent to a subordinating conjunction 'when; if; as'.

I could go on, but I think the point has been sufficiently made. What it boils down to is that much more rigor needs to be exercised from both the semantic and the phonetic point of view if one is to put the comparison of Old Chinese and Tibeto-Burman with other languages or even with each other on a solid basis. Otherwise one's view will be obscured by large amounts of irrelevant material.

As I said at the outset and as those who have followed my own work will be aware, I am not against the idea of long-range comparison in principle. I have long been developing a hypothesis about structural parallels and, ultimately, a genetic relationship between Sino-Tibetan and Indo-European. It is not based, however, on accumulating a long list of possibly cognate "look-alikes" at a superficial level, something that has been attempted many times in the past without carrying conviction, from Edkins (1871) and Schlegel (1872) in the nineteenth century to Shafer (1963, 1965) and Ulving (1968) in more recent times. Rather, it depends on a deep analysis of the underlying phonological structure of both language families and aims at the reconstruction of the same, rather simple, phonological system in both cases. It then emerges that the two language families also share basic morphological processes, including the morpheme **a* 'in' both as independent particle and as an infix, prefix and suffix, **š* (= **xʲ*) as a prefix with an intensifying force, **-n* and **-t* as stative (?) and punctual (?) suffixes, etc. The hypothesis also leads to the identifying of possible lexical correspondences (Pulleyblank forthcoming). If viable, however, these will have to meet the same kind of rigorous tests of regularity that are applicable in the traditional comparative method.

My starting point is that in both cases the vowel system as such was based on a single, two-way, opposition between the low vowel /a/ and a non-low schwa vowel /ə/, the upper corners of the universal *i-a-u* triangle being represented by the semivowels (glides) /j/and /w/ as independent phonemes and secondary articulations of other consonants. Unfortunately, this has been ruled out by most people as either impossible or as so unusual as to be not an acceptable model for a reconstructed language. There are such languages, however. The most familiar (or notorious) examples are in the Northwest Caucasian family, but such analyses of Modern Mandarin have been current for more than half a century. In the case of Indo-European, the treatment of /i/ and /u/ as allophones of the corresponding semivowels is part of standard theory. The remaining controversy is over how to account for the ablaut relationships between the three non-high vowel **e*, **a*, and **o*. It has been shown that the vowel traditionally reconstructed as **a* is in many, perhaps all, cases of secondary origin, reflecting a lost "laryngeal". The other two vowels, **e* and **o*, are linked by ablaut relationships. According to the traditional account, **e* alternates with the

reduced or zero grade when not accented, while **o* alternates with **e* in morphological categories and is assumed to be derived from it, though the supposed phonological conditions have never been satisfactorily explained. My hypothesis (Pulleyblank 1965a, 1993b) is that **e* is actually a central schwa vowel which appears and disappears according to rules of accent and syllabification and, in the absence of traditional **a*, **o* is a low vowel /a/ that has a parallel morphological function to /a/ in Kabardian, converting "extrovert" forms to "introvert" forms (Kuipers 1960). The same kind of morphological role is played by **a* in Sino-Tibetan (Pulleyblank 1965b, 1973, 1989). Sharing the same type of vowel system would not, of course, imply genetic relationship. Though comparatively rare, vertical vowel systems of the Northwest Caucasian or Mandarin type have also been attested in Australia and New Guinea. The semantic role of the one underlying vowel phoneme, /a/, is harder to explain as arising by chance. Moreover, in both language families, I believe I can identify the role of the morpheme **a* 'in' in other contexts — as a prefix and suffix and as an independent preposition. There are, I believe, also other correspondences in morphological processes at a deep level.

Unfortunately, my hypothesis has been not so much refuted as ignored. Though my two-vowel analysis of Old Chinese has been called "significant" (Ting 1975:32) and "simple and symmetrical, even elegant from an abstract point of view" (Baxter 1992: 813), both these reject it *a priori* without further discussion of its merits in solving problems of the internal development of the language. The proposed morphological explanation of the Indo-European **e/*o* ablaut has also been rejected without serious discussion on *a priori* grounds, in spite of the fact that no traditional explanation in terms of conditioned sound change has ever been found satisfying (Szemerényi 1967, Schmidt 1993, Lehmann 1993). It may be that I have not achieved sufficient clarity in my attempts to show that there is a common "introvert/extrovert" semantic opposition in contrasts such as noun *versus* verb, intransitive *versus* transitive, middle *versus* active voice, perfect (action completed) *versus* imperfect (action going on), absolutive *versus* relational particles, etc. In the absence of a purely phonetic explanation, however, the parallel with Kabardian as described by Kuipers seems to provide a typological model that ought at least to be seriously considered. The corollary that, if accepted, the semantic parallel associated with the same morpheme, **a*, seems to imply a common origin is, perhaps, an alarming prospect that many people are unwilling to contemplate. Unfortunately, I do not know how Starostin accounts for extrovert/introvert morphology of Northwest Caucasian in his reconstruction of the Proto-North-Caucasian. He does not mention it in his comparisons with Yenisei and Sino-Tibetan.

As for Chinese, the numerous alternations between **ə* and **a* among words that are otherwise clearly related in both sound and meaning are also not discussed., let alone explained, by such recent investigators as Li, Ting, Starostin, and Baxter, no doubt because even to acknowledge their existence would be to accept the possibility that my two-vowel analysis of the vowel system might be correct. This is by no means my only complaint against these scholars, who seem to me to be too limited and traditional in linguistic theory and methodology as well as in the types of evidence they employ (Pulleyblank 1992, 1993a). But these are questions that do not directly concern the present discussion.

References

Allen, W. Sidney. 1953. "Relationship in comparative linguistics," *Transactions of the Philological Society*, 52-108.

Baxter, William H. III. 1992. *A Handbook of Old Chinese Phonology*. Berlin and New York: Mouton de Gruyter.

Benedict, Paul K. 1942. "Thai, Kadai, and Indonesian: A new alignment in Southern Asia," *American Anthropologist*, n.s. 44: 576-601.

———— 1972. *Sino-Tibetan: a Conspectus*. Contributing editor: James A. Matisoff. Cambridge: Cambridge University Press.

Bloomfield, Leonard. 1935. *Language*. Revised edition. London: Allen and Unwin.

Edkins, Joseph. 1871. *China's Place in Philology. An Attempt to Show that the Languages of Europe and Asia have a Common Origin*. London: Trübner.

Kuipers, A.H. 1960. *Phoneme and Morpheme in Kabardian*. s'Gravenhage: Mouton.

Lehmann, W.P. 1993. "Comment on "The Typology of Indo-European" by Edwin G. Pulleyblank", *Journal of Indo-European Studies* 21:119-121.

Li, Fang-kuei. 1974-75. "Studies on Archaic Chinese," *Monumenta Serica* 31, pp. 219-287.

Miller, Roy Andrew. 1988. "The Sino-Tibetan Hypothesis," *Bulletin of the Institute of History and Philology, Academia Sinica* 59:509-540.

Pulleyblank, Edwin G. 1965a. "Close/open ablaut in Sino-Tibetan." In G. B. Millner and E. J. A.Henderson, eds. *Indo-Pacific Studies* (= *Lingua* 14), pp. 230-240.

———— 1965b. "The Indo-European Vowel System and the Qualitative Ablaut." *Word* 21: 86-101.

———— 1973a. "Some new hypotheses concerning word families in Chinese," *Journal of Chinese Linguistics* 1:111-125.

———— 1988. "Jo chih ho → nai ho," *Bulletin of the Institute of History and Philology, Academia Sinica* 59:339-351.

———— 1989. "Ablaut and initial voicing in Old Chinese morphology: **a* as an infix and prefix," *Proceedings of the Second International Conference on Sinology*. Section on Linguistics and Paleography. Taipei: Academia Sinica. pp.1-21.

———— 1992. "How do we reconstruct Old Chinese?" *Journal of the American Oriental Society* 112:365-382.

———— 1993a. "Old Chinese phonology: a review article." *Journal of Chinese Linguistics* 21:337-380 (Review of Baxter 1992).

———— 1993b. "The typology of Indo-European." *Journal of Indo-European Studies* 21:63-118.

———— 1995. "The Role of Glottal Stop in Old Chinese Phonology". Paper for the 4th International Conference on Chinese Linguistics. Madison, Wisconsin, June 26-30, 1995.

———— forthcoming a. "The Historical Relationships of Chinese." To appear in the Proceedings of a Symposium on the Ancestry of Chinese held following the 3rd International Conference on Chinese Linguistics. City Polytechnic of Hong Kong, July 14-16, 1994.

———— forthcoming b. "Response to Starostin's proposal for a connection between Sino-Tibetan, Yeniseian and North Caucasian." To appear in the Proceedings of a Symposium on the Ancestry of Chinese held following the 3rd International Conference on Chinese Linguistics. City Polytechnic of Hong Kong. July 14-16, 1994.

Ruhlen, Merritt. 1994. *On the Origin of Languages*. Stanford: Stanford University Press.

Sagart, Laurent. 1994. "Some Remarks on the Ancestry of Chinese." Paper for the Symposium on the Ancestry of Chinese, following the 3rd International Conference on Chinese Linguistics, City Polytechnic of Hong Kong, July 14-16, 1995.

Schlegel, Gustaaf. 1872. *Sinico-aryaca*. Batavia.

Schmidt, Karl Horst. 1993. "Comments on Edwin G. Pulleyblank: The Typology of Indo-European." *Journal of Indo-European Studies* 21:123-133.

Shafer, Robert. 1963. "Eurasial". *Orbis* 12:19-44.

———— 1965. "The Eurasial Linguistic Superfamily," *Anthropos* 60:445-68.

Starostin, S. A. 1988. "Methodology of Long-range Comparison", in Vitaly Shevoroshkin, ed., *Nostratic, Dene-Caucasian, Austric and Amerind*. Bochum: Brockmeyer. pp 75-79.

———— 1989. *Rekonstrutsiya drevnekitajskoj fonologicheskoj sistemy*. Moscow: Nauka.

Starostin, S. A. and Merritt Ruhlen. 1994. "Proto-Yeniseian Reconstructions, with Extra-Yeniseian Comparisons". In Ruhlen 1994.

Szemerényi, Oswald. 1967. "The new look of Indo-European: reconstruction and typology". *Phonetica* 17:65-99.

Ting, Pang-hsin. 1975. *Chinese Phonology of the Wei-Chin Period: Reconstruction of the Finals as Reflected in Poetry*. Taipei: Institute of History and Philology, Academia Sinica.

Ulving, Tor. 1968. "Indo-European elements in Chinese," *Anthropos* 63/64:943-951.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

COMMENTS ON R. L. TRASK'S CRITIQUE: IS BASQUE AN ISOLATE?

Merritt Ruhlen
Palo Alto, California

[Editor's Note: Dr. Ruhlen is an original long ranger, a founding member of ASLIP, and a frequent author of books on long range comparisons. With John Bengtson, he has proposed several sets of global etymologies. His current address is 4335 Cesano Court, Palo Alto, CA 94306. Tel. 415-948-3248]

If the central myth of twentieth-century historical linguistics has been the claim that the Indo-European family has no known genetic connections with any other family, the first corollary of this myth has been the notion that the Basque language has no visible genetics links with any other language. Both of these myths have had vigorous critics throughout the twentieth century, but their defenders — primarily Indo-Europeanists — have been no less vociferous in their defense. In recent years the first myth — the splendid isolation of Indo-European — has begun to crumble under the weight of the evidence offered by the Nostraticists (especially Vladislav Illič-Svityč, Aaron Dolgopolsky, and Allan R. Bomhard) and by Joseph Greenberg in a series of articles and a forthcoming book on the Eurasiatic family.

In the paper here under consideration (Trask 1995), an eminent Basque scholar has examined all of the alleged evidence connecting Basque with other languages and families and has concluded that "the evidence so painstakingly assembled for relating Basque to the other 'Dene-Caucasian' languages amounts to precisely zero. Basque remains as isolated as it ever was." I will examine this conclusion in the remainder of this paper.

First, however, we should note that the title of Trask's paper, "Basque and Dene-Caucasian", is somewhat misleading inasmuch as Trask later admits that he has "ignored the data from Sino-Tibetan and Na-Dene." Trask does not explain why he has chosen to ignore two of Dene-Caucasian's six branches, but in so doing he has clearly forgone the possibility of evaluating the Dene-Caucasian hypothesis, which would require showing that the other four branches of Dene-Caucasian — Basque, Caucasian, Burushaski, Yeniseian — are closer to Sino-Tibetan and Na-Dene than they are to, say, Khoisan and Australian. (I ignore Sumerian, whose position in Dene-Caucasian is more controversial.) Rather than truly evaluating the Dene-Caucasian hypothesis, Trask ignores the work of Starostin (1991) on Sino-Caucasian and that of Nikolaev (1991) on Caucasian-Na-Dene. His paper is thus limited to an examination of proposed similarities between Basque and Caucasian, Burushaski, or Yeniseian. Proposed similarities with Sino-Tibetan or Na-Dene — even when they strengthen etymologies among the four groups Trask considers — are silently ignored. This is a curious practice, all the more so because the proposed Sino-Tibetan and Na-Dene evidence is given in the same articles that Trask used for the rest of his paper.

To see the consequences of Trask's only considering two-thirds of Dene-Caucasian, it is instructive to consider etymology 230. Bengtson's proposed connection of Basque *odol* 'blood' with Burushaski *del* 'oil, contents of an egg' seems phonologically plausible, but semantically only possible. Thus a linguist evaluating the evidence given by Trask would no doubt retain a good degree of skepticism regarding the proposed etymology. However, were this same linguist given the evidence from Na-Dene (contained in Bengtson's article but ignored by Trask) — Eyak *deɬ* 'blood', Proto-Athabaskan **deɬ* 'blood', Carrier *del-el* 'blood', Galice *daɬ* 'blood', Navajo *diɬ* — he would be more likely to conclude that the entire etymology is almost certainly valid.

Nor can Trask's paper be considered an evaluation of Bengtson's proposed subgroup Macro-Caucasian, which would unite Basque, Caucasian, and Burushaski. The validity of this subgroup would require taking into account all of Dene-Caucasian, something Trask has not done. (Note, however, that etymologies 57-68 given below appear to be exclusively shared innovations that directly support Bengtson's Macro-Caucasian hypothesis.) What Trask's paper is then is an appraisal of the evidence connecting Basque with three other families.

Though Trask concludes that the evidence connecting Basque to other languages is "zero", he also admits that there is nothing apparently wrong — at least from the Basque side — in roughly half of the 317 proposed etymologies. These 150 left-over etymologies, without any apparent defects, Trask dismisses as "vague resemblances . . . [that do not] constitute evidence for anything." Let us take a look at some of these "vague

resemblances" so that we may better judge for ourselves. In each of the following etymologies Trask either has no serious objection to the Basque data, or his objections are so far-fetched that I have ignored them. For example, Trask's only objection to Bengtson's proposed triconsonantal root for 'dung' — Basque *korotz* ~ *gorotz*, Proto-Caucasian **k'urč'V*, Burushaski *γurAš* — is that the Basque form is supposed to be a borrowing of Latin CROCEA 'saffron-colored', which, according to Trask, "is phonologically impeccable." The reader should take special note of the fact that *all* the data from Sino-Tibetan and Na-Dene given in the etymologies below were omitted by Trask.

The following abbreviations are used: PC: Proto-Caucasian, PWC: Proto-West Caucasian, PEC: Proto-East Caucasian, Bur: Burushaski, Bur(W): Werchikwar dialect of Burushaski[1], PY: Proto-Yeniseian, PST: Proto-Sino-Tibetan, PTB: Proto-Tibeto-Burman, OC: Old Chinese, PND: Proto-Na-Dene, PEA: Proto-Eyak-Athabaskan, PA: Proto-Athabaskan. The number of Trask's etymology is given in brackets at the end of each etymology. Non-linguists should be aware that what is written *z* and *x* in Basque are really kinds of *s*'s, the first similar to a normal English *s*, the second similar to English *sh* ([š] - Ed.) as in 'she'. Furthermore, *c*, *tˢ*, and *ts* in these etymologies probably represent the same sound, the final *-ts* in 'ca*ts*'.

Grammatical Similarities

1. Basque *no-r* 'who', *no-n* 'where', *no-iz* 'when', *no-la* 'how'; PC **nV* (interrogative), Agul *na* 'who', Abkhaz *-an(ə)-* 'when'; Bur (W) *ana* 'where'; PY **ʔan-* 'who', Ket *an-a* 'who'; PST **nāi(ŋ)* (interrogative), OC *nāj* 'how, what', Tibetan *na* 'when'. [229]

2. Basque *ze-r* 'what', *ze-rin* 'which', *ze-nbat* 'how much'; PC **s̄a* (interrogative pronoun base), Avar *s̄u-n* 'what', Hunzib *si-n* 'what', *se-j-* 'who, which'; PY **ʔas-* ~ **sV-* 'what', Kot *še-na* 'what'; PST **su* 'who', Tibetan *su* 'who'; ?Navajo *-(i)š* 'what'. [302]

3. Basque *ni* 'I'; PC **nɨ* 'I'; PA **nu-hani* 'we', Chepewyan *nu-hni*, Navajo *ni-hi*. [227]

4. Basque *-t* ~ *-da-* 'I'; PC **t-* 'we', Adyghe *te* 'we', Chamalal *itɬi* 'we (inclusive)'; Haida *-īt^{l}'* 'we', PA **-ī'd* 'we', Hupa *-it-* 'we', Navajo *-ii(d)-* 'we'.[2] [266]

5. Pre-Basque **-ga* 'thou (masculine)'; PC **Gu* 'thou'; Bur *gu-* 'thou'; PY **kV* 'thou'; PST **kʷVj* 'thou'. [145, 191]

6. Basque *d-* (third-person marker); PC **d-* (third-person animate singular marker). [80]

7. Basque *zu* 'you (pl.)'; PC **źwV* 'you (pl.)', Abkhaz *šʷa-* 'you (pl)', Adyghe *śʷa* 'you (pl.)'; PEC **źwV* 'you (pl.)'. [312]

8. Basque *ez* 'no, not'; PC *ćə* (negative); Bur *ačho* 'not yet'. [103]

9. Basque *-k* (plural); Abkhaz *-kʷa* (plural); Bur *-ko(ŋ)* (plural); PND **-qai(ʔ)* (plural), Tlingit *-x'*, Haida *-χa*, Eyak *-qɛiyu*, PA **-k(e)*, Navajo *-ke*. [189]

---

[1] Editor's note: cf. Tiffou's remarks about the names of the dialects of Burushaski, in his comments on Trask's paper. By his suggestion, Werchikwar ought now be called Yasin.

[2] Editor's note: We do not know the proposed phonetics of the reconstructed PA form. It might be **-ī'd*, making it close to [ɗ], an imploded [d].

10. Basque *-tzu* (plural of indefinites); Abkhaz *-c^{w}a* (plural), PEC **-cwV* (plural). [278]

11. Basque *-ez-* (adjective); Bur *īs-* (adjective); PY **-se* (adjective). See etymology No. 60 below for Basque and Burushaski examples. [24]

12. Basque *-n* (past tense); Abkhaz *-n* (past tense), PEC **-na* (past tense). [225]

13. Basque *ra-* (causative); Abkhaz *ar-* (causative). [249]

14. Basque *-z* (instrumental); PEC **-s(ē)* (instrumental); Ket *-aś* (instrumental), Kott *-s(e)* (instrumental). [291]

15. Basque *-k* (ergative); PC **-k'V* (ablative, instrumental); Bur *-Ak ~ -ek* (instrumental). [190]

16. Basque *-i* (dative); PC **-i* (dative); Bur *-e* (genitive, oblique). [155]

17. Basque *-la* (manner); Abkhaz *-la* (manner). [202]

18. Basque *-n* (locative); Abkhaz *-n* (locative), PEC **-nɨ* (locative). [224]

Lexical Similarities

19. Basque *agor* 'dry'; PC **ʔĭ-G^{w}Vr-* 'dry, dry up', Lezghi *q̄urq̄ur*, Lak *q'arq'-*; Bur *qAqər* 'very old and wrinkled'; PY **qVr$_1$* 'dry'; PST **qār* 'dry'; Haida *g'a ~ k'a* 'dry', Tlingit *q'uan* 'dry salmon', PEA **g^{w}aN(H)* 'dry', PA *Gɑŋ* 'dry', Ingalik *-Gəŋ* 'dry', Chepewyan *-gan* 'dry', Navajo *-gan* 'dry'. [7] Note the sound correspondence Proto-Dene-Caucasian **-r* = Proto-Na-Dene **-n*, also seen in etymologies 26 and 50 below. Additional evidence for this sound correspondence, first noted by Nikolaev (1991:46) for Na-Dene and Caucasian, is given in Ruhlen (1995).

20. Basque *bil-du* 'collect, gather';Bur *bil* 'full', *bilbinʌs* 'fill up' (intransitive); PTB **bliŋ* 'full, fill', Rong *blen-* 'full, fill', Digaro *bloŋ* 'full, fill', Mikir *pleŋ* 'full, complete', Tshungli *ben* 'full, fill'; PA **-wən* 'fill with liquid', Mattole *-bil ~ bīŋ ~ bīʔn* 'to be full or become full', Navajo *-bīɬ ~ -bin* 'fill'. [71]

21. Basque *egun* 'day'; PC **ʔwɨ̄-Gɨ̄nʌ* 'day'; Bur *gōn* 'dawn', Bur(W) *gun* 'dawn', *gūn-c* 'day'; PY **χōŋ* 'day(time)'; OC **k^{w}āŋ* 'light'; Haida *kūŋ* 'sun', Tlingit *Gugan* 'sun', *seiGun* 'tomorrow', Eyak *ǰah* 'day(light)', Tanaina *ǰanih* 'day(light)'. [83]

22. Basque *gose* 'hungry'; PC **gāši* 'hunger', Lak *k̄aši*, Agul *gaš*, Dargwa *guš*; Tibeto-Burman: Hruso *k^{h}usśĭ* 'hungry'; Athabaskan: Galice *gas* 'become hungry'. [120]

23. Basque *izan* 'be'; PC **ʔos^{w}V* 'sit'; Bur *ōs* 'set'; PY **hAs-* 'be'; Haida *is* 'be'. [175]

24. Basque *hats* 'paw, leg (of an animal), finger (of a human); PC **k^{w}ač'e* 'paw'; Bur *qAṣ* 'length from the elbow to fingertips' ; PY *kiʔs* 'foot'; Eyak *-kaša* 'fingernail', PA **-kеč'* 'claw'.[134]

25. Basque *igel* 'frog'; PC **q̄'^{w}VrV-q̄'V* 'frog'; Bur *gʊr-qʊc* 'frog'; PY **xiʔr-* 'frog'; Eyak *čiaɬq* 'frog'; Athabaskan: Hupa *č^{w}'aɬ* 'frog', Navajo *č'aɬ* 'frog'. [158]

26. Basque *izar* 'star'; PC **Ha-d̄ᶻʷarʕi* 'star'; Bur *lA-čār* 'morning star'; PTB **tśar* 'star, sun, moon'; PA **swenʔ ~ *səŋ$_2$ʔ* 'star', Upper Kuskokwim *ṣunʔ*, Tanaina *sən*, Tanana *sənʔ*, Tolowa *sənʔ*, Galice *sanʔ*, Navajo *sõ*. [176]

27. Basque *izen* 'name'; PC **- d̄ᶻʷěr-* 'name', Adyghe *c'a*, Chechen *c'e*, Udi *c'i*, Andi *c̄'er*; Bur *sen-As* 'named'; Sino-Tibetan: Tibetan *m-can* 'name'; PND **ciaʔ* 'name', Tlingit *sen* 'to name' (v.), *sa* 'name' (n.), Athabaskan: Ahtena *zä*, Sarsi *í-zìʔ*, Galice *-šiʔ*, Navajo *yí-ží*. [178]

28. Basque *karats* 'bitter'; PC **q'elʔV ~ *q'eq'elʔV* 'bitter'; Bur *γāqAy* 'bitter', Bur (W) *qʌqā-m* 'bitter, sour'; PY **qʌqʌr* 'bitter, bile', Yug *xʌxilʲ-aŋ*, Kott *ogar* 'bile, gall', Ket *qʌlʲ* 'bile, gall'; PST **ghāH* 'bitter', OC **qāʔ* 'bitter', Tibetan *kʰa-ba*, Garo *kʰa*, Burmese *kʰa*; ?Haida *k'aywaɬ* 'be sour'. [197]

29. Basque *odol* 'blood'; Burushaski *del* 'oil, contents of an egg'; Eyak *deɬ* 'blood'; PA **deɬ* 'blood', Carrier *del-el* 'blood', Galice *daɬ* 'blood', Navajo *diɬ* 'blood'. [230]

30. Basque *sagu* 'mouse'; PC **cārgwɨ̄* 'weasel, marten, mouse'[3]; PY **saʔqa* 'squirrel'; PST **srăiŋ*, OC **srjeŋ* 'weasel', Burmese *hraṅ* 'squirrel'; PND **ts'ɑlk(ʔ)* 'squirrel', Haida *čolgɪ* 'gopher', *gaɬ-tˢ'akʷ* 'squirrel', Tlingit *tˢaɬk'* 'squirrel', Eyak *tˢaɬk'* 'squirrel', PA **tˢələx ~ *tsəlxʸɪ* 'squirrel', Carrier *tˢolq-* 'chipmunk'. [253]

31. Basque *soin* 'shoulder, body, torso'; PC **čwěnʔV* 'game, wild animal'; Bur *šAn* 'limbs, all the parts of the body'; OC *śjěn* 'body', PTB **sya* 'flesh, meat'; Eyak *-ceʔ* 'flesh', PA **-cənʔ* 'flesh, meat', Beaver *tsûnʔ* 'meat', Carrier *ceŋ*, Galice *sanʔ* 'meat'. [262]

32. Basque *sunda* 'stench, rank smell'; PC **sūn-t'V* 'to smell; odor'[4]; Bur *ṣûŋ* (*ɛtʌs*) 'to smell'; PTB **suŋ* 'smell', Karen *suŋ* 'odor', Tibetan *b-suŋ* 'smell sweet'; PND **čan* 'stench', Haida *sánjuu* 'smell', Tlingit *čan* 'stench', Eyak *čãh* 'stink', PA **-čən* 'smell', Galice *šan* 'smell', Kato *čən* 'smell'. [264]

33. Basque *tu* 'spit, saliva'; PEC **tuj* 'spit, saliva'; Bur *thū* 'spit'; PY **duq* 'saliva'; PTB **(m)-tuk* 'saliva, spit'; PND **tuχ* 'spit', Tlingit *tuχ* 'to spit', Eyak *tux* 'saliva', Kato *šek'* 'saliva'. [269]

34. Basque *txorru* 'root (of a hair)'; PC **č'aHVrV* 'hair'; Bur *-čhōqur-* 'forelock'; Tlingit *šuχawō* 'hair'. [276]

35. Basque *(h)ur* 'water'; PC **ɦwiri* 'lake'; Bur *hūr* 'conduit for water', *hʊr(-iginAs)* 'stream'; PY **xur$_1$* 'water'; PST **ɣu-s* 'moisture'. [284, 159]

36. Basque *sasi* 'bramble'; PC **cāce* 'burr'; Bur *čhAš* 'thorn'; PY **seʔs* 'larch'. [257]

37. Basque *zuri* 'white': PC **hʷö-č'ʷörV* 'gray, yellow'; PY **suɾ́-* 'yellow'. [316]

---

[3] Editor's note: In fairness to Bengtson, one mentions that in Catford (1991:264) there is a rare list of supporting forms (for a PC starred form); some 8 reconstructions and 10 specific language citations, under 'weasel / mouse' #73. It has to be noted, also, that Bengtson has been generally handicapped by the lack of such forms.

[4] Editor's note: Again in fairness to Bengtson, it was argued in MT-1 and/or MT-2 that the PC form (then PEC) is really **sūnt'* — In fairness to Trask, we note that the original cut was arbitrary.

38. Basque *milika* 'to lick'; PC **mělc'ĭ* 'tongue'; Bur *-melč* 'jaw'; PST **mlaj* 'tongue', Longla *-mili*, Ao Naga *-moli*, Katchin *mli-ak* 'to lick'. [218]

39. Basque *galtzar* 'side of the body'; PC **ʔā-gʷāɬʷV* 'side'; PY **χol-* 'cheek'; PST **qālH* 'back, cheek'; Eyak *l-quhɬ* 'cheek'. [106]

40. Basque *ukondo* 'elbow'; PC **q'ʷV̨ntV* 'elbow, knee'; PY **gid* 'elbow, joint'; Eyak *Guhd* 'elbow', Athabaskan: Sarsi *gūd* 'elbow, knee'. [283]

41. Basque *eri* 'finger'; Bur *-rīŋ* 'hand'; Tibeto-Burman: Mikir *ri* 'hand'; Tamang *(nā)-ri* 'arm'. [93]

42. Basque *aztal* 'calf of the leg, heel'; PC **t'Hālq'ʷV* 'part of leg', Proto-Lezghian **t'elq'Iʷ* 'shin, ankle', Proto-Tzez **tɨq'ʷV* 'sole of foot'; Bur *-l-tAlt-er* 'foreleg (of a quadruped)'; PA **-tAdˡ'* 'heel', Chepewyan *-tal* 'heel', Navajo *-tal* 'heel'. [34]

43. Basque *esku-barne* 'palm of the hand' (*esku* = 'hand'); PC **bar-gwa* 'palm of the hand'; Tibetan *s-par* 'palm', Tsangla *par* 'palmful'. [97]

44. Basque *hertze* 'intestine'; PC **p̄ĕHĕrtˡ'V* 'intestine'; Bur *wArk* 'fat on the big intestine of animals; kind of sausage'; PY **pɨʔɨlˀ* 'intestine'; PST **Pik* 'intestines', PTB **pik* 'intestines', Mikir *phek* 'intestines'; PND **-wadʔ* 'stomach', PA **-wədʔ* ~ **-wət'* 'belly', Mattole *-baʔɬ*, Wailaki *-bit'*, Navajo *-bid'*. [143]

45. Basque *haitz* 'stone, rock'; Bur *čiš* 'mountain, hill'; PY **čɨʔs* 'stone', Yug *čɨʔs*, Ket *tɨʔś*, Kott *šīš*, Assan *šiš*; Haida *t'īs* 'stone, rock sticking out of water', Tlingit *ēč'* 'peeble, round stone', Eyak *čīš* 'beach sand, gravel', PA **ce* 'stone', Mattole *cē*, Chasta Costa *se*, Navajo *ce*. [127]

46. Basque *ihintz* 'dew'; PC **xănfiɨ* 'water', Proto-Dargwa **xin* 'water', Tsakhur *x́an* 'water'; PND **Gan-i* ~ **han-i* ~ **Xan-i* 'water, river', Haida *ʕantˡ* 'fresh water, creek', Tlingit *hin* ~ *hain* 'fresh water, river', Eyak *ʔã* 'river', PA **xanʔe* 'river', Tutchone *xanʔe*, Kutchin *han*, Hupa *hanʔ*. [159]

47. Basque *agure* 'old man'; Bur *gūro* 'grayish'; OC **grjij* 'old', Tibetan *b-gre-s* 'to grow old', Hruso *mo-kʰro* 'old man', Dimasa *găra-sa* 'old man'. [8]

48. Basque *atso* 'old woman'; PC **d̄ᶻʷVjV* 'female', Avar *c̄'o* 'woman'; Bur *-n-co* 'paternal aunt'; PND **ču* 'grandmother', Haida *joo-n* 'mother-in-law', Tlingit *ča-n* 'mother-in-law', Eyak *-čū* 'grandmother', PA **-čo-ŋ* 'mother-in-law', **čwə-nə* 'grandmother', Hupa *čʷoo* 'grandmother', Navajo *-čo* 'grandmother'. [36]

49. Basque *azkoin* 'badger' (< **hartz-konH* 'bear'?); PY **kūñ* 'wolverine'; Eyak *kənaʔs* 'wolverine'. [40]. This etymology indicates that Michael Krauss was in fact correct in his dispute with Jürgen Pinnow (Krauss 1976:339-40) concerning the Eyak form given above; it is not related to the Haida, Tlingit, and Athabaskan forms cited by Pinnow, and the initial *kə-* is not a prefix but simply part of the root, as argued by Krauss.

50. Basque *egur* 'firewood'; PC **gɔ̄rV* 'stick, pole'; Tlingit *gun* 'firewood', Eyak *də-kīh* 'stick', PA **-kən* 'stick'. [84]

51. Basque *hats* 'breath'; Chechen *ħožu* 'odor', Batsbi *ħaič'* 'odor'; Bur *hiš* 'breath'; PA **-γīk'* ~ **-γīč'ʷ* 'to breathe' (< Pre-Proto-Athabaskan **-xīk'ʷ*). [133]

52. Basque *adar* 'horn, branch'; PC **tɬVrV* 'horn'; Bur *-ltūr* 'horn'; ?Eyak *-dəleh* 'horn', PA **-deʔ* 'horn'. [3]

53. Basque *bigar* 'tomorrow'; PC **bəgə* 'morning', Khinalug *pāga* 'tomorrow', Proto-Dagestan **pəkV* 'morning, dawn'; PY **pʌk-* 'morning'. [69]

54. Basque *zilbor* 'navel'; Tsakhur *ǰil* 'navel'; PY **tɨl-* 'navel'. [307]

55. Basque *lepo-zaki* 'back of neck, nape' (= 'neck-bone' according to Azkue); Haida *χil tˢ'ak'ī* 'back of neck' (= neck back), Tlingit *s'aq* 'bone'. [295]

56. Basque *gogor* 'hard' (< **gorgor*); PC **gʷērV* 'stone'; Bur *γoro* 'stone'; ?Haida *qwā* 'rock'. [117]

57. Basque *korotz ~ gorotz* 'dung'; PC **k'urč'V* 'dung'; Bur *γurAš* 'dung'. [199]

58. Basque *mutur* 'snout, muzzle'; PC **mərtɬ'V* 'nose, beak'; Bur *-multur* 'nostril'. [233]

59. Basque *begi-nini(-ko)* 'pupil of the eye' (*begi* = eye); PC **nănV* 'pupil of the eye'; Bur *nAna* 'eyeball'. [52]

60. Basque *ar* 'male', *arr-ez-ko* 'masculine'; PC **Hir-k'ʷV* 'man'; Bur *hīr* 'man, male', *hir-īs-ki* 'man's, men's'. [24]

61. Basque *pirpirin* 'butterfly'; PC **părVpăɬV* 'butterfly, moth', **pɨrV* 'to fly'; Bur *pfer-An* 'moth'. [244]

62. Basque *zikiro* 'castrated goat'; PC **c'ɨ̄kV* 'goat, kid'; Bur *chīgir* 'nanny goat'. [306]

63. Basque *gozo* 'sweet'; PWC **q'aş'ʷə* 'sweet'; Bur *gaš-ar-um* 'salt-sweet'. [121]

64. Basque *beso* 'arm'; PC **bəcV* 'paw, part of arm or leg', Tsez *besi* 'fist', Khinalug *beši* 'fist', Tabassaran *bac* 'paw'. [64]

65. Basque *hilindi* 'charcoal, something scorched, firebrand'; PC **ɬwind* 'firewood'. [148]

66. Basque *ibar* 'valley, water meadow'; Bur *bAr* 'valley, ravine'. [156]

67. Basque *intzaur* 'walnut' (<**intza-hur* '?-nut'); Bur *xʊnzʊr* 'kernal of walnut'. [165]

68. Basque *sahats* 'willow tree'; Bur *šAsk* 'willow tree'. [254]

What are we to make of the above resemblances, which represent a fraction of the proposed evidence for Dene-Caucasian? According to Trask, "we are clearly looking at nothing but an assembly of chance resemblances between languages which have apparently been selected on some *a priori* basis as grist for comparison." If Dene-Caucasian is just an *a priori* arbitrary combination of random families, it is certainly an audacious one. To claim that Basque in Europe, Chinese in Asia, and Navajo in the American Southwest are more closely related historically to one another than to other languages in closer geographical propinquity to each is indeed an audacious claim. Yet the Dene-Caucasian proposal — in various shapes in the work of different scholars — has attracted the support of Alfredo Trombetti, Edward Sapir, Morris Swadesh, Karl Bouda, Sergei Starostin, Sergei Nikolaev, John Bengtson, Vyačeslav Čirikba, Václav Blažek, and others during this century. Trask would have us believe that all these scholars have merely deluded themselves into seeing historical relationships where none exist.

How are we to resolve the status of Basque? Numerous scholars have connected it with a certain set of languages, yet Trask maintains that all the evidence amounts to zero. However, what Trask is implicitly claiming in his dismissal of the Bongo-Bongo approach is that he could find equally cogent similarities between Basque and any other five linguistic families. After all, Dene-Caucasian is just an *a priori* concoction of six arbitrarily selected families. If this is true, then Trask should be able to come up with equally cogent evidence connecting Basque with five other families. To my knowledge no one has ever attempted this. Why is it then that a number of scholars during this century have investigated the relationships of Basque essentially in terms of the Dene-Caucasian hypothesis, while no one — so far — has come up with what Trask considers an equally plausible hypothesis, the Austro-Basque hypothesis, which would join Basque with Australian, Khoisan, Gilyak, Algonquian, and Quechuan. This would indeed be a daring and audacious proposal, but if everything is *a priori*, who cares?

Who cares should be linguists interested in the history of Basque. If part of the legacy of Nostratic/Eurasiatic studies has been the realization that certain aspects of Indo-European can only be understood in the larger context, the lesson for Basque, amply demonstrated in the Dene-Caucasian etymologies, is that many aspects of Basque can only be understood in terms of the Dene-Caucasian hypothesis. The bizarre semantic shifts that Trask invokes, the mysterious borrowings from unidentified Romance sources, and the outright rejection of any etymology that violates his preconceptions about Pre-Basque phonology — all of these expedients are used by Trask to do away with putative evidence, half of which nonetheless withstands all attack.

The final etymology that I cited above involves a comparison between Basque and Burushaski that I find instructive: Basque *sahats* 'willow tree' and Burushaski *šAsk* 'willow tree'. Though it involves just two isolated languages spoken thousands of miles apart, it draws our attention first of all by its semantic identity. Trask may claim that the Dene-Caucasian proposal involves loose semantics, but in the enormous universe of semantic space a matching of 'willow tree' with 'willow tree' is about as precise as one can get. Furthermore it just so happens that the word in both languages involves two sibilants and another consonant. What is the probability that the word for 'willow tree' would consist of two sibilants and another consonant by accident? I suspect there are no other languages which have such a word.

Let us then assume that these two words probably are related, rather than dismissing them out of hand, as Trask does. Is there any way to reconcile the slight phonetic differences between the two forms? Trask reminds us that there are four dialect variants of this word in Basque — *sahats* ~ *sagats* ~ *sarats* ~ *saats* — and that "this type of variation points unmistakably to a lost intervocalic consonant and hence an original **saCats*." But what consonant, precisely, was *C*? On the basis of the Basque evidence I would hypothesize that the original form in Pre-Basque as **sakats*, even though this is not an attested dialect form. In some dialects the intervocalic *-k-* was fricativized to *-h-* (perhaps via an intermediate stage *-x-*), in others it was voiced to *-g-*, in others it was rhoticized to *-r-*, and in some it was lost entirely. These are all natural phonetic changes.

What if the Pre-Burushaski word was identical with the Pre-Basque word. Is there any way to go from Pre-Burushaski *sakats* to Burushaski *šAsk*? I would suggest the most likely route is: **sakats* > **sakas* > **saks* > **sask* > *šask* > *šAsk*. These changes too represent natural phonetic shifts, abundantly attested in the world's languages. Furthermore, if our analysis is correct, then Proto-Yeniseian **ʔɔkse* 'wood' is just one sibilant dissimilation away from the hypothetical Pre-Burushaski form **saks*.

The lesson is clear: progress in Basque historical linguistics will only come when Basque specialists begin to explore Basque within the context of the Dene-Caucasian hypothesis. The glory days of splendid isolation — both for Indo-European and for Basque — are, with the twentieth century, drawing to a close.

References

Krauss, Michael E. 1976. "Na-Dene," in *Native Languages of the Americas*, Vol.1, ed. by Thomas A. Sebeok, New York, Plenum, pp. 283-358.

Nikolaev, Sergei. 1991. "Sino-Caucasian Languages in America," in *Dene-Sino-Caucasian Languages*, ed. by Vitaly Shevoroshkin, Bochum, Brockmeyer, pp. 42-66.

Ruhlen, Merritt. 1995. "A Dene-Caucasian Sound Correspondence," ms.

Starostin, Sergei A. 1991. "On the Hypothesis of a Genetic Connection Between the Sino-Tibetan Languages and the Yeniseian and North-Caucasian Languages," in *Dene-Sino-Caucasian Languages*, ed. by Vitaly Shevoroshkin, Bochum, Brockmeyer, pp. 12-41.

Trask, R.L. 1995. "Basque and Dene-Caucasian: A Critique from the Basque Side," in this issue of *Mother Tongue.*

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

ON BASQUE AND OTHER SINO-CAUCASIAN (DENE-CAUCASIAN) LANGUAGES

Vitalij Shevoroshkin
University of Michigan

[Editor's Note: Vitalij Shevoroshkin is Professor of Linguistics, University of Michigan. He is an expert on the Anatolian branch of Indo-European, especially Carian, and more generally involved in Nostratic, Dene-Caucasic, and Amerind studies. In the editor's opinion, Vitalij is the leading representive of the Muscovite school of long range comparisons in the United States, although he might not accept that label. His address is 1133 Michigan Avenue, Ann Arbor, MI 48104. Telephone (312) 663-9462. His text has been lightly edited.]

In independent comparisons of Basque with the North Caucasian languages (V. Čirikba, J. Bengtson, et al.), solid evidence of a genetic relationship of these languages has been presented. This evidence was strengthened by recent work in Russia which has led to an important reconstruction of the Proto-North Caucasian language as presented now by S. Starostin in the form of a solid book.

As usual, we deal with a fair amount of misunderstanding when discussing work on remote relationship of languages: R. Trask's impression that Proto-North Caucasian (PNC) was a language with 180 consonants is just one example of such misinterpretations of available data. Several years ago, Starostin and Nikolaev did indeed reconstruct too many consonants for PNC, but this was done deliberately, in order to keep apart some "difficult" stems which should represent a certain proto-stem, not yet reconstructed at that point. Later, it became clear that the alleged stems did not show archaic features that made them different from each other; they rather contained innovations in certain languages due to internal prosodic processes in the latter. All this allowed the authors to drastically reduce the number of PNC consonants, excluding those which equaled sounds — present as innovations — in certain subgroups of NC languages.

It is wrong to consider Bengtson, Ruhlen, and myself as architects of the Dene-Caucasian (DC) idea: the priority belongs to S. Nikolaev, who showed that NC (consisting of East- and West-NC), ST (Sino-Tibetan) and Yeniseian, united as Sino-Caucasian (SC) languages by Starostin, are also related to Na Dene / Athapascan and some other languages in North America.[1]

As I see it now, Nikolaev's idea about "Almosan" languages (Salishan and Wakashan) being a part of Dene-Caucasian (Nikolaev's term!) is very sound; moreover, archaic Salishan data considerably improve Starostin's reconstruction of SC, which is based only on three languages.[2] I tried to illustrate this idea on several occasions: cf., for instance, my introductory remarks to the book *Proto-Languages and Proto-Cultures*, edited by Vitalij Shevoroshkin (1990, Bochum, N. Brockmeyer), p. 9: "... as soon as I started to compare Salishan with Sino-Caucasian, I saw that practically all stablest roots ... show clear matching between Salishan and SC (mostly between Salishan and NC)". I am going to cite just one example (#11 'two') from my list:

Salishan (Moses-Columbian) *t-q'aw's* (*t-* is a Proto-Salishan prefix) : SC = NC **t'-ʔq'wE* (*t'-* is a prefix).

Starostin correctly compared this SC root/word with Nostratic **tuʔV*, as reconstructed by A. Dolgopolsky in the early sixties. Note that the Kartvelian word for 'pair' is a borrowing from NC, not an inherited lexeme.

---

[1]See Nikolaev, "Eyak-Athapascan — North Caucasian Sound Correspondences", in Vitalij Shevoroshkin, ed., *Reconstructing Languages and Cultures*, 1989. Bochum, N. Brockmeyer. Also "Sino-Caucasian Languages in America", in Vitalij Shevoroshkin, ed., *Dene-Sino-Caucasian Languages*, 1991. Bochum, N. Brockmeyer. [Editor's Note: This note was in the text above.]

[2]See over 200 SC roots, compared with Nostratic, in Sergei Starostin. 1989. "Nostratic and Sino-Caucasian", in Vitalij Shevoroshkin, ed., *Explorations in Language Macrofamilies*. Bochum, N. Brockmeyer.

I am deeply convinced that any comparisons of Basque with other DC languages, to be plausible in all cases, will have to include data from all (proto-)languages of this phylum (where cognates are present). Still, comparisons by Bengtson and Čirikba show many sets which are "real". (By the way, when speaking to me lately in Leiden, Čirikba told me that he was going to eliminate some 40% of his Basque-to-NC sets as imprecise, replacing them with dozens of more plausible comparisons.)

As usual in languages, the most important comparisons between Basque and NC (or SC, for that matter) should include the stablest words from Dolgopolsky's list (based on the hierarchy of stability of word/roots in question; primarily, roots which are not borrowed from language to language); data after Bengtson and Starostin:

1. Basque *ni* 'I' : NC **nɨ* (: Sal[ishan] **nɨ*);
2. Basque *hi* 'thou' : NC, SC **Gu* (: Sal. : Sq[uamish] *ʔa-x*^{w} = related SC **ʔV-G*);
3. Basque root *no-* 'who, when' : NC **nV* (interrogative)
4. Basque *ez, ze* (dial.) 'not' : NC **ć(ʾ)ə* 'not';
5. Basque *u-khondo* 'elbow' : NC **qʾHwəntV* 'elbow, knee' (*'joint');
6. Basque *bortz* 'five, hand' : NC **b̲ɔrcV* 'paw';
7. Basque *hama-r* 'ten' : NC **xamħ̄V* 'handful'
8. Basque root *gal-* 'side' : NC **ʔā-G*^{w}*ā*^{ɬw}*V* (Consonantism as in Basque *gune* 'place' : NC **G*^{w}*ī̲HnhV* 'village, hut', etc.);
9. Basque *be-larri* 'ear' : NC **l̲eHl̲e* 'ear';
10. Basque *i/ulhe* 'wool' : NC **ə̄lXHV* 'wool' [Not among the most stable!]

And so on; this kind of comparison is quite preliminary: one should perform an internal reconstruction of Basque first, and check NC data in Starostin's recent *North Caucasian Etymological Dictionary*; one should also compare PNC / PSC roots not directly with Basque and its dialects but with reconstructed Proto-Basque, from which phonemes should be taken for a comparison with NC (or SC, for that matter) phonemes. I am sure, such a comparison would strongly corroborate a very sound hypothesis about the NC[3] character of Basque.

One may prefer a different approach: one may ignore the solid comparative-historical evidence and compare Basque *bartz* 'louse' with *var-* 'worm' in Romance languages, etc., though 'louse', 'nit', 'flea' are among the the most stable words and should be compared with such. To ignore linguistic evidence is counterproductive: people start arguing and stop working.

[3]Dolgopolsky's term for NC is Hurri-Caucasian.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

ABOUT TRASK: BASQUE AND DENE-CAUCASIAN: A CRITIQUE FROM THE BASQUE SIDE

Étienne Tiffou
Université de Montréal

Address: Faculté des arts et des sciences, Département de linguistique et de traduction, Université de Montréal, C.P. 6128, succursale Centre-ville, Montréal, Québec H3C 3J7, CANADA

[Editorial preface: Dr. Tiffou as a linguist has been primarily interested in Burushaski, including its history and classification. He has also contributed to field research on Burushaski and Yasin (Werchikwar). His strong secondary interest in Basque is also relevant here.]

In a long study which could be the subject matter of a book, R. L. Trask critically examines the arguments which have been given to legitimize the classification of the Basque language in the Dene-Caucasian family. He scrutinizes particularly J. Bengtson's, V. Blažek's, M. Ruhlen's, V. Čirikba's, and S. Starostin's theories. In fact, his review, if it is valid, tends to demonstrate that this language cannot be included in this important linguistic branch; however, to do so, it is sufficient to question if the Basque really belongs to a subfamily: the Macro-Caucasian one. This subfamily would include the Caucasian languages, Basque, and Burushaski. Therefore, Trask only attempts in his study to disqualify the assumption according to which the Basque language has to be considered as belonging to this subgroup.

The introduction offers an historical survey of the problem. Then, Trask, after he has reminded us of the important number of Basque words whose the origin is Latin or Romance, gives seven fundamental laws of Basque phonology. Next, this scholar gives a close look at 317 terms which are supposed to share some similarities with other languages. Trask's conclusion discusses, on one side, the lessons to be drawn from his analysis, while on the other side, he critically examines some morphological parallels and calls into question the relevance of phonological charts matching Basque, Caucasian, and Burushaski phonemes. The results of this study are clear: it is impossible to determine any linguistic family, whatever that may be, to which Basque may belong. The arguments put forth in an attempt to prove links with other languages are not at all convincing. According to Trask, we are dealing, not with valid proposals, but with a collection of badly matched facts.

In his analysis, Trask notes that he will not criticize the data from languages other than Basque — he accepts the cited forms from other languages as they are given. However, he scrutinizes the Basque words on which the Macro-Caucasian hypothesis is based. It is only occasionally that he questions the proposed similarities without appealing to the internal analysis of Basque. More often, his comments are based on his exemplary knowledge of this language, but he also finds a very important help in the outstanding works of L. Michelena, chiefly in the *Phonética historica vasca* (1976) of this scholar. The statement devoted to the Basque loan words (pp. 5 sq.) is essential. The Basque language borrowed a lot of words from Latin and the Romance languages. These loan words have been integrated in such a way that it is often difficult to detect their genuine origin. Therefore, we must be completely sure, when we want to propose parallels between Basque words and foreign words, and so to demonstrate the links between all these languages, that the Basque words under consideration are not loan words. The phonological rules which follow, introduce indirectly and briefly the internal reconstruction of Basque. A meticulous observation of every studied word induces one, more often than not, to draw a negative conclusion.

Before I draw my own conclusion on the problem concerning the classification of Basque, I would like to express a point of view matching Trask's analysis. To do so, my study will be based on another language. I will, at least I hope, meet Trask's wishes when he writes (p. 104): "Given the authors' track record as regards Basque, is it likely that a Burushaski specialist, if any exist, would find the Burushaski data to be error free?" I considered myself always as one of the linguists specializing in Burushaski; we can count also among them H. Berger (1974 and forthcoming), E. Bashir (1985), A. Frémont (1982), B. Tikkanen (1991), S. Willson, and my colleagues Y. Morin (1989) et R. Patry (Tiffou and Patry 1996).

Anyway, I would like to make a wish — is it possible from now on to give up calling Yasin Burushaski

"Werchikwar"? This term ["Werchikwar" —editor] is used for this language by Chitrali people; in the same way, Shin people call Hunza Burushaski "Khajuna", but never, as far as I know, do linguists use this word. Yasin people and Hunza people use the word *burúšaski* 'the language of the Burushos' or the word *mišáski* 'our language', to call their own language. On another side, Hunza Burushaski and Yasin Burushaski represent two dialects of the same language, while Hunza Burushaski and Nagar Burushaski are two subdialects very close to each other. Hence, there is no reason to consider, as one did too often, that Yasin Burushaski is a degenerated Burushaski dialect or a subdialect of Hunza Burushaski (see Müller-Stellrecht, 1973:51), and we have to speak respectively of Hunza Burushaski (HB) and of Yasin Burushaski (YB). Actually, Hunzukuts tend now to consider that YB is more reliable than HB, because big changes occurred in Hunza, and they think that contact with people speaking foreign languages is corrupting HB. This is, of course, not true. Anyway, I have to say that I was very surprised by the corpus studied by Trask. Of the 131 references to Burushaski, only two concern YB. This seems to me particularly strange because we have very good information on this dialect (Zarubin 1927; Lorimer 1935/I:422-452, 1935/II: 399-415, 1938:394-417, and 1962; Berger 1974; Morin and Tiffo, 1989; Tiffou and Pesot 1989). Finally, I am convinced that on many points and particularly from a morphological point of view, YB is more conservative than HB. Hence, it is wrong not to take this Burushaski dialect into consideration in historical linguistic studies.

Anyway, starting from Burushaski, it is interesting to corroborate or to confirm Trask's analysis, because of the 317 entries he considers, as I already mentioned, Burushaski is involved 131 times, namely, in a little bit more than 41% of the entries. I will not study every case, but only some significant examples.

[4.] Bq. *adar* 'horn, branch', Bur. *-ltōr, tur*

These two words have the same meaning in the two languages. However, it is difficult to establish a connection between them at a phonological level. In fact, the correct notation is *-ltúr* (the *u* is not long and is stressed); the form *tur* is used when the prefixed pronoun (which marks inalienable possession) is at 1st sg. pr. If so, how is it possible to justify the assumed similarity with Bq. *adar*?

[10.] Bq. *ahizpa* 'sister (of a woman)'; Bur. *-ACo*, *-ACu* 'sister (woman speaking), brother (man speaking)'.

I have a preliminary remark: we have to give up Lorimer's notation, for this one is phonetic and not phonemic. It is better to adopt Berger's notation. This is followed by 'Allamah Nasir al-Din Hunzai (see *Dīwān-i Naṣīrī*), Tikkanen (1991), and myself (1989, 1993), with negligible changes from time to time. Therefore, for example, we would have to write-°*Co* / -°*Cu* instead of *-ACo*, *-Acu*, where *A* is a transliteration of Lorimer's a̲. I have to remind you that *-/-'* is the notation of type I prefix (for example, YB *guyéca* 'I see you'); -̲° is the notation of type II (YB *gópai* 'beside you'); -" is the notation of type III (YB *gō'mala* 'you are afraid'). It is this notation I will use in the following comments.

I don't see phonological similarities between the Burushaski and the Basque words; moreover, the expression of brotherhood is based on four words in Basque and only three in Burushaski. As a matter of fact, we have in Basque *ahizpa* for a sister's sister, *arreba* for a brother's sister, *anai* for a brother's brother, and *neba* for a sister's brother. On the other hand, -°*Co* / -°*Cu* designates in Burushaski brothers and sisters of the same sex. One man's sister is called in YB *-yást* and *yás* in HB; a woman's brother is called in YB *-húlés* and in HB *-úlus* respectively. According to all these remarks, we can consider that the comparison between Basque and Burushaski is not very convincing.

[24.] Bq. *ar* 'male'; Bur. *hiir*.

In spite of Klimov and Edel'man (1970:25), I never elicited the long vowel *ī*; I am sure that *hir* has never had a long vowel in YB; and what about the relation of Bq. *a* and Bur. *i/ī*?

[29.] Bq. *arraultza* 'egg'; Bur. *i-rič* 'kidney'.

The Burushaski word would have to be noted as *-'riD*. I elicited it only in YB. I don't see the relation between eggs and kidneys. This comparison seems to me very suspect because Burushaski has a very ancient word for 'egg' — *ṭiŋán*.

[43.] Bq. *bai* 'yes'; Bur. *A-wa* 'yes'.

There is little to say about these words, but I am really surprised. There is a formal similarity between the Basque word 'yes' and the Burushaski word for 'no' — *be*, reinforced form *béia*. If the cut given by Trask is plausible, as the stressing of the Burushaski word (*awá*) and the attestation in HB of a particle *wa* which indicates surprise or reinforce the assertion may argue in favour of this interpretation, however, the similarity between Bq. *b* Bur. *w* requires an explanation. Why, in the more convincing comparison given in [156], is there a correspondence between a Basque *b* and a Burushaski *b* (Bq. *ibar* 'vallJe'; Bur. *bar* 'vallJe')? We cannot pretend that initial Bur. *b* becomes *w* in another position, for there is not in Burushaski an assertive particle *ba*.

[62.] Bq. *bero* 'hot, warm'; Bur. *babárum* 'hot, warm'.

This parallel is whimsical. There is a phonological similarity between these two words, but the meaning is very different. The authors have been misled by an English homonymy. 'Hot' in this language means 'hot, warm' and 'spicy, highly seasoned'. *Bero* has the first meaning and *babárum* the second one.

[88.] Bq. *emakume* 'woman'; Bur. *quma ~ kuma* 'concubine'.

We have to reject this parallel, for the word *quma* is a loan word borrowed from Khowar. Furthermore, the variant *kumá* is suspect. The meaning 'concubine' has been given by Lorimer, but in his book *Das Yasin Burushaski*, Berger proposes the translation 'morganatic wife'. If Berger is right, this meaning has to be preferred, because — even though HB is near to Shina — YB is near to Khowar.

[93.] Bq. *eri* 'finger'; Bur. *-riiŋ* 'hand'.

In my opinion, the comparison between these two words is questionable for semantic reasons. HB *-ríiŋ* matches the Basque word better, but the YB one *-rén* seems to me older. Anyway, why would the Burushaski word for 'hand' correspond to the Basque word for 'finger'? This would be very surprising, because this language has an ancient word for 'finger', to wit, HB *-°miṣ*, YB *-°meṣ*? The same remark applies to [96] where the Basque word *esku* 'hand' is compared with the Burushaski word *hesk* 'wrist, back of the hand'.

[104.] Bq. *ezti* 'honey'; Bur. *mačhi* 'honey'.

YB has the form, but without the aspirate, i.e., *maṭí*. However ,the penultimate phoneme is not palatal, but retroflex. This mistake occurs frequently (see for instance 50 and 104). We know that the opposition between a retroflex and a palatal is phonological in Burushaski. This confusion may give rise to big errors.

[136.] Bq. *haur* 'self'; Bur. *-khar* 'self'.

There is a lot to say about this entry. As Trask notes, Basque *haur* does not mean 'self', nor does *-khar* in Burushaski. In this language, this word is used to express the reflexive; the corresponding word in Basque is *buru*. Cf. Bq. *bere burua hil du* 'he killed himself' and YB *akhár ésqana* 'I killed myself'. In Basque, the word *buru* means 'head' and is periphrastically used to mark reflexivity (the translation word by word of the Basque sentence is 'he killed his head'). The meaning of *-khar* is not clear, but, like *buru*, it is used periphrastically. This can be inferred from the resumption of this word with the type II verbal prefix, which has to be at the 3rd singular

(*ésqana*). Here, we have an important typological similarity between Basque and Burushaski, but this one is found in other languages (cf., for instance, the use of *jibun* in Japanese).

[155.] Bq. *-i - -ri* (dative); Bur. *-e* (genitive, general oblique).
The value of general oblique that the ending *-e* is supposed to have, is right. One can suppose there was formerly an absolutive case distinct from a morphologically characterized case. This one is well known with an ergative or genitive value. We can find it with a locative value — *Yásine* 'in Yasin'; but I could never elicit the dative value with this ending. If this case had been expressed by the general oblique, it probably became independent very early. Under these conditions, the matching with the Basque dative is strange. Why, on the other hand, should we suppose that a Basque *i* would correspond to a Burushaski *e*? In this language, *i* does not change in final position, e.g., *íŋi* 'before him'.

[159.] Bq. *i-hintz* 'dew'; Bur. *huu-š* 'wetness (of ground), moisture'.

The transcription in Burushaski is spurious; the vowel is not long, and the final consonant is retroflex, thus *huṣ*.

[167.] Bq. *ipini* 'put'; Bur. *-mAn-(As)* 'be, become'.

The comparison is phonologically more acceptable with the Basque variant *min-*, but I cannot accept the parallel between an intransitive and a transitive verb, for *-mán- / man-'* is always intransitive, and a transitive verb has never been derived in Burushaski from this one. The same comment can be made for [175].

[190.] Bq. *-k* 'ergative case-suffix'; Bur. *-Ak / -ek* 'instrumental'.

I have never seen the instrumental suffix, here mentioned, neither in HB, nor in YB.

[199.] Bq. *korotz* ; Bur. *γurAš* 'dung'.

The comparison is shaky. *γúraš* designates man's excrements and sometimes cow's dung. *Korotz* designates animal's excrements; the approximate corresponding word in Burushaski is *nayé*.

[217.] Bq. *matel* 'cheek'; Bur. *mAto* 'brains'.

The Burushaski word seems to be a loan word from Shina, whereas the word *bal* (in Burushaski) which has two meanings 'marrow' and 'brain', is the original Burushaski word.

[229.] Bq. *nor* 'who?' ; Bur. (W) *a-na* 'where, whither', *a-n-Um* 'whence'.

We can't draw any conclusion from the YB indefinite / interrogative / relative adverb *án-*, for it is difficult to analyze it clearly. However, *a-na* and *a-n-Um* have, for sure, to be reanalyzed as *án-a* and *án-um*, where it is easy to identify respectively the dative ending and the locative one.

[250.] Bq. *-ra-* 'allative case-suffix'; Bur. *-ar* 'allative'.

The Burushaski suffix has not only an allative value but also a dative one. Which is the older? It is not easy to tell. One has to observe that the proposed suffix is HB. In YB, the suffix is *-a / -γa*. This form fits better with the Basque ending.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

A COMMENT ON "BASQUE AND DENE-CAUCASIAN: A CRITIQUE FROM THE BASQUE SIDE" BY R. L.TRASK.

Roger Williams Wescott
16-A Heritage Crest
Southbury, CT 06488-1370

[Editor's Note: Roger Wescott is a founding father of ASLIP and one of the original long rangers. Indeed it was his recommended designation 'long ranger' for members of the Long Range Comparison Club that started our self-label. Roger is without doubt the most catholic of long rangers, with interests running into anthropology and comparative study of civilizations. He is Professor Emeritus from Drew University in New Jersey.]

Trask's Critique is sobering. One need not accept his negativism about long-range comparison to acknowledge the desirability of some of his recommendations. The first of these is that, in comparing Basque forms with forms from other languages, we should cite as many dialectal variants of each form as possible. The second is that, in reconstructing pre-Basque, we should, as far as possible, eliminate all of the many Latin and Romance loan-words in modern Basque. And the third is that, in the same reconstructive process, we should do our best to avoid violating the apparent phonotactic rules of pre-Latin Basque.

A fourth such recommendation might be that all long rangers dealing with Basque should consult what seems to be the most complete work on the linguistic history of Basque: *Sobre Historia de la Lengua vasca*, two volumes, authored by Luis Michelena and edited by J. A. Lakarra, Donostia-San Sebastián: Anejos del Anuario del Seminario de Filología Vasca: Julio de Urguijo: 10, 1988.

Trask's other cautions about the search for languages cognate with Basque, however, seem to me less daunting. He warns, for example, against arbitrary segmentation of forms. Where contemporary languages are concerned, familiarity with their morphology should make erroneous segmentation unlikely. But reconstructed languages are inevitably problematic, since no reconstruction is beyond question. And interpretation of reconstructed forms is, if anything, even more arguable. In Proto-Indo-European, for example, it is doubtful how such numerals as **k*^{w}*etwores* 'four' and **oktōu* 'eight', should be segmented and doubly doubtful whether the syllables beginning with *t* should be regarded as alternate forms of the numeral **dwōu*.

Trask regards the small consonantal inventory of 'native' Basque and the large consonant inventory of Caucasian languages as constituting a typological barrier to genealogical links between them. Typological discrepancies between clearly related languages, however, are so common that I cannot regard this obstacle as insurmountable.

If Bengton or other long rangers claimed that Basque is closely related to any other language or languages, Trask's insistence on systematic sound correspondences would be quite justified. But all regard the relationship as distant.[1] This distance makes it likely that phoneme alternations and sporadic variations will have blurred such consistent one-to-one correspondences as may once have existed between the ancestors of these languages.

As regards the Basque/Caucasian connection, Trask gives Michelena as his authority that "no evidence of any significance at all" has been presented in its favor. What constitutes "evidence", however, is a matter of opinion. And "significance" is, if anything, an even more subjective term. Such wording, while effective in expressing attitudes, does little to inform. Moreover, when Trask asserts that Michelena's conclusion "must be accepted by anyone who takes historical linguistics seriously," he clearly implies that long rangers are frivolous scholars. Innuendo of this kind, I think, begins to exceed the bounds of collegial civility.

Trask's characterizations on the last page of his article are still more emotive. There he writes, for example, of long rangers as comparing Basque "breathlessly" with languages geographically remote from the Pyrenees. And

[1] Editor's note: What is close and what is distant can be defined differently as between scholars. Bengtson has said that Basque is not so far from Caucasic and Burushaski, probably not much farther apart than Indo-European branches. Fleming thinks those are close or at least obvious.

he concludes that evidence relating Basque to those languages is "precisely zero". At this point, he clearly engages in hyperbole. For most of his article consistes of a list of Basque words whose cognation with "Dene-Caucasian" words he is at pains to refute in detail. Yet some words, such as Basque *bete* 'full' (related by Bengtson, Ruhlen, and Starostin to Burushaski *būt* 'much') are left unrefuted. In other words, even Trask admits that evidence for the thesis he rejects exceeds zero. What remains is to discover how far beyond zero it goes.

My own conclusion is that linguistic isolation is a relative rather than an absolute phenomenon. In terms of mutual intelligibility, every language is by definition an isolate. (Otherwise it would be a dialect rather than a language.) In terms of ultimate cognation, however, every language manifests — at least for linguistic monogeneticists — kinship with all other languages. The question then becomes one of determining not the existence but the extent of its isolation. And, for this, extensive on-going comparative work continues to be required.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

COMMENTS ON R. L. TRASK'S ARTICLE "BASQUE AND DENE-CAUCASIAN: A CRITIQUE FROM THE BASQUE SIDE"

Xabier Zabaltza
Yanguas y Miranda, 23
31002 Iruñea
Nafarroa Garaia
Euskal Herria, SPAIN

[Editorial preface: Dr. Zabaltza is a historian, especially of Basque areas, and a native speaker of Basque, especially dialects in northwestern Spain. He is also member of ASLIP and describes himself as a long ranger.]

When I received the article by R. L. Trask about my language, Basque, and its relations with Dene-Caucasian, the first thing I did was to look at the references in order to see if the best Bascologist who has ever existed, Luis Michelena, was cited. He was. The second thing I did was to scrutinize the bibliography to see if there were any articles written in Basque. There were.

I did not expect less of Dr. Trask, as I knew some of his works on the Basque language, which show a great knowledge of it. Unfortunately, this is not the case of all the authors who write about Euskara. I wonder what could be the value of an article about English literature by somebody who has never heard of Shakespeare or, even, who only knows a few words of the English language. Nothing, I suppose. So, the reader will be able to imagine what I feel when I see an article about my language, written by someone who has no idea who Luis Michelena was or is incapable of analyzing correctly the most simple word in Basconia's Mother Tongue.

When Dr. Fleming asked me to be a referee within our association, I answered that it would be an horour for me, but that I did not know if I would be of any real use, as a referee is supposed to refute the text (s)he is given and I was sure I would agree with most of Trask's arguments.

Therefore, the aim of these comments is to support Trask's thesis, according to which there is not yet any proof of the relationship between Basque and Na-Dene and Caucasian languages. I can say I agree with 99% of Trask's affirmations. In fact, I have to recognize that if I had written an article about this question I would have tried to do one very similar to Trask's. On these lines, I will have to resign myself to comment on the more or less one per cent of affirmations I do not agree with or into which I wish to introduce some nuances. Some of them are, in my opinion, mere *lapsus calami*, and I consider that Dr. Trask will not have any problem in correcting them in the definitive version of his work.

Basque, like all the languages on Earth, has evolved throughout the centuries. A Basque speaker of today would have at least as many problems understanding another one from two thousand years ago as a modern Spanish speaker would have understanding Cicero. However, we do not call Cicero's tongue "Old Spanish", but "Latin". On the other hand, the only words we have to denominate the ancestor of the present-day Basque language (for example, that of two centuries ago) are "ancient", "old" or "proto-" Basque. The word Basque always appears, and we tend to consider this language as homogenous throughout its long history.

Like most languages too, Basque has not only had a "natural" evolution. At least since the XVIIIth century (and probably since the XVIth), great purist efforts have been made to "extirpate" words derived from Spanish, French, and Latin. Many neologisms were created, and some dialectal words were spread quite artificially to substitute more common ones which were clear loans from Romance. I think Trask has not taken into account these "purist" efforts, which explain the success of words such as *apez* 'priest', *ulertu* 'understand', *uhin* 'vague', or *adin* (in the meaning of 'age'), instead of *abade*, *konprenitu*, *olatu*, or *edade* respectively. These successful words are anyway very difficult to recognize. Only with the help of specialized dictionaries (Michelena, 1987- and Sarasola, 1984-1995) will the scholar manage to find out something worthwhile.

As Michelena himself used to say, only comparable things can be compared. This obvious remark is easily forgotten in comparisons between Basque and Caucasian languages. On the one hand, we have a single language, Basque, with six dialects: Guipuzcoan, Biscayan, High Navarrese, Low Navarrese, Labourdin, and Souletin (Roncalese disappeared earlier this century, and Salacenco and Aezcoan, both mentioned by Trask, are usually

considered as subdialects of Low Navarrese). On the other hand, according to Ruhlen (1987:324), the Caucasian languages were believed to constitute one family of 38members. As the reader will probably realize, the likelihood of finding cognates ("random similarities" is Trask's expression) between one of these 38 languages and Basque leans towards 100%. I suspect the results would be exactly the same with every other language family which has 38 members. The only way to avoid this chaos, needless to say, is to use proto-languages as instrument of comparisons and, without a doubt, it is Michelena's task that would have to be used as a base for Basque reconstructions.

More recently (Ruhlen, 1987:379), the Caucasian family has been split up in two different families, Northern Caucasian (34 languages) and Southern Caucasian or Kartvelian (4 languages). Only the Northern Caucasian languages are believed to take part in the Dene-Caucasian phylum, in which Basque is included too. This division leaves out all the work of some very competent Basque scholars (such as Kintana, 1981) who, little by little, but with a solid base (in my opinion), have made advances in the investigation of the relationships between Basque and Georgian, which now is considered not (North) Caucasian, but Kartvelian (South Caucasian).

A temptation of every comparativist is to "forget" the loans. For instance, there are some words which, seemingly, could be regarded as possible Basque and Kartvelian cognates (Kintana, 1981:263). One of them is Basque *orein* and Georgian *iremi* 'deer'. The Basque word derives from **oleni* and the Georgian one from **ilen-*. But both can be perfectly explained by means of Indo-European **eln-*. In the same way, the Georgian *zeti* and the Spanish *aceite* 'oil' are obviously loanwords from Arabic *zéit* and no Hispano-Caucasian root is needed.

Comments on Trask's List of Basque Words

abets 'voice'. It must be a mistake. I think two Basque words have been mixed up: *abots* 'voice' and *abesti* 'song'. Both of them are neologisms and both contain *abo* or *aho* 'mouth', but only the second of them was coined by the Basque nationalist writer Sabino Arana. The first, *abots*, (in today's Basque, *ahots*), was invented by Manuel Larramendi in 1745. What Sabino Arana did was to spread it. (This is not a unique case: neologisms coined by Larramendi and other authors, such as *ikastola* 'school', *askatasun* 'freedom", *guda* 'war', and *gudari* 'soldier', are usually and wrongly considered as having been created by Arana). Anyway, the considerations on some Basque 'fantastic neologisms' and Sabino Arana's 'absurd speculations' are not linguistic, but ideological.

adin 'age'. The ordinary meanings of this word are 'understanding, judgment' (to relate to *aditu* 'hear, understand') and 'contemporary'. Only in some parts of Basconia did it mean 'age'. It began to replace the common word for 'age' (which is *edade*, obviously from Spanish *edad*) in the middle of the XVIIIth century, on account of the purist efforts of the previously mentioned Larramendi. The second element of *adin* is the suffix *din*, of verbal origin, which appears in several Basque words: *gordin* 'raw, crude', *urdin* 'cloudy, turbid', *berdin* 'similar, alike'. See Michelena 1970: 295.

asto 'donkey'. Trask puts in question the etymology of this word given by Azkue and Michelena (**hartz-to*, something like 'little bear'), adducing the Aquitanian divine name *Astoilunno deo*. This problem could be solved if the Aquitanian *asto* was related to *aste* 'week' (Michelena, 1954a:433). Even in modern English, the days of the week take their names from gods.

atal 'segment, fragment, portion'. An added problem with this word is that it cannot have been documented before Larramendi's dictionary (1745). One perfect equivalent of *atal* is the Navarrese *ziztor*, from which the name *txistorra*, the typical Basque sausage, derives.

axuri 'lamb'. I agree with Trask's etymology, as white is the color of sheep in Basque tradition.

azkoin 'badger'. It has been compared with Latin *taxone(m)* (Michelena, 1964:72). The loss of the initial consonant might have favored a folk etymology which related it to *hartz* 'bear' (this could explain the Souletin form *harzku*). If this is true, the second element, interpreted as **koin* and used to establish the Proto-World **kuan* 'dog', would be just the second syllable of the Spanish word *tejón* 'badger'.

-ba (in kinship terms). At the beginning of the XIXth century, the Basque author Pablo Pedro de Astarloa interpreted this root of the kinship terms as *aba* (without any doubt, in order to compare it with Hebrew *av* 'father'). It served the Basque nationalist leader Sabino Arana as a way of coining neologisms such as *aberri* 'fatherland', *abenda* 'race', and *abizen* 'family name'. On the other side, the Basque writer Arturo Campion at the beginning of this century deduced *eba* from *-ba* (influenced by the name of the first woman according to the Bible, which in Basque is Eba), to interpret it as 'woman, female'.

barakurkuilo 'snail'. Better *barakuilo*. The second element is the same as in Spanish *caracol* 'snail' (there are possible intermediaries, such as *karakoil* and *barakoil*). It might be a folk etymology, influenced by *bare* 'spleen, slug'.

barrabil 'testicle'. For the form, *barret* and *barrasaket* 'sterile' must be loans from Gallo-Romance.

beginini 'pupil of the eye'. In Spanish, *niña* has also both meanings, 'child' and 'pupil'. The Basque form *nini* is, in all probability, borrowed from Spanish.

belarri 'ear'. Michelena (1976:221) relates this word to *beha* 'to pay attention, to listen'.

bete 'full'. It is commonly believed to be related to *bat* (from *bade*, Michelena, 1976:134), 'one' (Perurena, 1993:40). See expressions like *ordubete* 'one hour' (etymologically 'full hour'), *egunbete* 'one day', *astebete* 'one week', *hilabete* 'one month', *urtebete* 'one year'.

bide 'road, way'. There is the possibility of it being a Latin loan (*via*), although the intermediary is not clear. See the name of *Bidasoa*, the river which separates the Spanish Basque province of Guipuzcoa from the French Basque province of Labourd, where the first element might be either Latin *via ad* or Basque *bide*. (The second element is clear, Oiasso, modern Oiartzun, one of the most important Basque towns during Roman times, Michelena, 1956:141).

bihar. This word does not mean 'yesterday', as Trask without any doubt knows, but 'tomorrow'. It must be a mistake.

egun 'day'. The *n* is probably a fossilized inessive (Michelena, 1976:138, note 1). *Egun* still means 'today' in the northern dialects, because *gaur*, 'today' in the other dialects, is clearly derived from *gau* 'night'. It has nothing to do with *hego* 'south'.

emakume 'woman'. One problem with this word is that it does not appear in Basque literature before 1627. The second element is *ume* 'child', from **unbe*, related to the Aquitanian word *ombe* (Michelena, 1979:346)

ero 'crazy, stupid'. Related to *erho* 'to kill', the causative of *jo* 'to hit, to beat' (Michelena, 1976:526).

esne 'milk'. From **es / ezende* (Michelena, 1976:163).

ez 'no, not'. From **eze*, which explains the common *ez* and the old Biscayan *ze*.

ezti 'honey'. Related to the name of the bee, which in Basque is *erle* (from * *ezle*, Kintana, 1981:264).

galdu 'to lose'. From **kal-* (Kintana, 1981:265).

gau 'night'. From *gau* plus *hau(r)* 'this' derives the adverb *gaur* 'today'.

gibel 'liver'. The second element is *bel(tz)*, which in modern Basque means 'black', but whose original meaning is 'dark'. The same suffix *bel* appears in several Basque words, such as *ubel* 'livid, purplish', *goibel* 'dark sky', *orbel* 'fallen leaf', *ezpel* 'box' (Perurena, 1992:55).

giltzurrin 'kidney'. The second element could be *urdin* (in modern Basque 'blue', but initially 'cloudy, turbid', related to *ur* 'water'). See Perurena, 1992:162.

hauts 'dust, powder'. Related to *hautsi* 'to break'.

hegal 'wing'. Related to *hegi* 'ridge', as both involve the idea of 'edge, corner'. See Michelena, 1949:450-451.

hi 'you, thou'. According to Martinet, an ancient glottalized (*kh*) could convert itself into *k, h,* or *Ø* at the beginning of the word and into *k* (masculine mark) at the end (Txillardegi, 1980:30).

ibar 'water-meadow or valley'. It is related to *ibai* 'river'. Tovar (1959:160) mentions the correspondence between these two words, but also between *bizkar* 'back' and *bizkai* 'hillock' and *hamar* 'ten' and *amai* 'end' (the last one not accepted by Michelena). On the other side, given the geographical distribution of Basconia, there is no doubt that *ibar* is related to *ipar* 'north' (Michelena, 1976:267, note 11).

***ika* 'one'. The origin of this false suffix is very interesting. It was invented by the Souletin canon Emmanuel Intxauspe in the second half of the XIXth century. He, in fact, "reconstructed" the form ***eka*, from the dialectal variety *hameka*, in order to compare it with Sanskrit *éka-s* and Hebrew *ehad* 'one'. His ***eka* might have been the common word for 'one' for all of Mankind before Babel (Perurena, 1993:225). The Basque nationalist author Sabino Arana even created some neologisms derived from this "word", which, have not been successful. In fact, *hamaika* 'eleven' derives from *hama-* 'ten' plus **bed* 'one' plus the suffix *-ka* (Michelena, 1976:496).

According to Trask, *hamaika*, apart from meaning 'eleven' is also the word used throughout the Basque country to denote an indefinitely large number, rather like English 'zillion'. This is absolutely true, but this value of *hamaika* is modern, from the middle of the XIXth century. Sentences like *nereak eta bost esan nizkion* 'I told him/her off' and *bost axola zait* 'I am not interested in it at all' prove that it was the number *bost* 'five' that had the meaning of 'zillion'. See Perurena, 1993:146, 225.

intzigar 'hard frost'. It has nothing to do with *aintzira* 'swamp, lake' or *aintzika* 'reservoir'.

itzal 'shade, shadow'. It is related to *itzali* 'to put out, to extinguish'.

izten 'awl'. Better *ezten*. Its first meaning is 'sting'.

***kala* 'castle' (?). I absolutely agree with Trask that the Aquitanian *kala* for 'castle' never existed (even if it had existed it should not be used in comparisons, as it is a cultural word, in Swadesh's terminology) Anyway, I find too rash Trask's affirmation according to which, Aquitanian or Basque "was never spoken as far south as Calahorra in Roman times". One of the most gripping debates, which has not been solved yet in my country, is about the frontiers of the Basque language in ancient times. As Trask knows, Calagurris was a Vascon city, although this does not necessarily mean that Basque was spoken there. The ancient name of Alfaro (in the Spanish province of Logroño, near Calahorra) was *Graccurris* and prior to it, *Ilurcis*. Both of them seem to contain the Basque element *hiri* 'town' (from **il-i*). According to Corominas (1972:250, note 15), Basque was spoken in the mountains of Soria (very much to the south of modern Basconia) until the Roman times, while Celtic languages were spoken in its lowlands.[1]

[1] Editor's note. For the readers' benefit we looked the modern names up and located them on maps. The area they are contained in is often listed on maps as *Cordillera Iberica* or the mountains between Old Castille and Aragon. This is

lagun 'companion'. I do not know if this word can be translated by English 'friend'. It means 'companion', but 'person, individual, inhabitant' too. The Basque for 'friend' is *adiskide*. There is the expression *denak lagun, adiskide guti* 'one must be companion of everybody, but friends are always few'.

maguri 'strawberry'. The common form for 'strawberry' is *marrubi*, which obviously derives from Latin *marrubium*.

-n (locative case suffix). In Basque, like in Indo-European, the difference between ablative and genitive is secondary. The *-n* of the Basque inessive is a simple variant of genitive *-ren* (Michelena, 1976:504).

ni 'I'. Trask shows the existence of two different roots for the first person singular *n* (as in *naiz* 'I am') and *t* (as in *dut* 'I have'). According to him, "This suggests an ancient stem-alternation in the Basque pronoun, and once again ... this is just the sort of puzzle on which we would expect a valid comparison to shed some light." The existence of *d* in the extinct Roncalese dialect (*dud* instead of *dut* 'I have', and *dakid* instead of *dakit* 'I know') shows that originally both were occlusives (nasal *n* and oral *d*). According to Martinet, this might confirm that the original pronoun of the first person singular was a prenasalized consonant /**nd*/, which might have evolved towards an *n* in the initial position and towards a *d* in the final position (Txillardegi, 1980:30).[2]

on 'good'. According to Trask, "The Aquitanian evidence suggests that the earliest form of the adjective might have been **bon*, with the frequent loss of *b* before *o*." The possibility of being related to Latin *bonus* was not totally discarded by Michelena (1954a:435).

or 'dog'. Better *hor*. It is usually derived from *hori* 'yellow' (in modern Basque). The names of colors are related to animals in this language. *zuri* 'white' is the color of sheep (*aziendazuri)*, *gorri* 'red' that of cattle (*abelgorri*) and *beltz* 'black', that of pigs (*aziendabeltz*). See Perurena, 1992:17.

oskol 'shell, peel, bark of a tree'. It has nothing to do with *azkazal* 'nail' (from *atz* 'finger' plus *azal* 'skin').

sabel 'stomach, belly'. The second element is *bel(tz)*, 'dark, black', like in *gibel* 'liver' (Perurena, 1992:55).

sahats 'willow'. Mujika (1982:168, note 464) prefers a Romance etymology (Latin *salix*). Michelena (1950a:462) does not share this opinion.

samin 'bitter'. The second element is *min* 'hot, peppery'.

-t 'I' (ending of verbs). See the remarks under *ni* 'I' above.

---

part of the area which McCall and Fleming, following Anderson, considered indubitably Basque proper, as opposed to Aquitanian or Iberian or Tartessian.

[2] Editor's note. This for the non-linguists. In many phyla, e.g., Indo-European and Afrasian, there are two first person singular pronouns, often with one in the focal role (e.g., English 'It is I', French 'c'est moi') and the other usually restricted to bound forms or contexts (e.g., English 'm-y, th-y' but also 'a-m'. Americans and British dialects tend to say 'It's me.') The two pronoun forms may or may not be linked in some ancient underlying form. Omotic, for example, has **n* and **t* which appear not to be linked. Yet the 'ego' and 'me' forms may be linked in Proto-Indo-European, with something like **egho-m* proposed in some contexts.

uhain 'wave'. The most common words for 'wave' are *olatu* (obviously from Spanish *ola*) and *baga* (from French *vague*). *uhin* (better than *uhain*) is a dialect word which has succeeded because of linguistic purism. I do not think it has anything to do with *ur* 'water' plus *gain* 'top'. It is a clear loan from Latin *unda* which means exactly the same.

zazpi 'seven'. The best etymology continues to be that of Michelena: *borzaz bi*, literally, 'two besides of five'. (Michelena, 1972:298, note 14).

zilar 'silver'. The intermediary between the Indo-European root and Basque *zilar* might have been Celtiberian *silabur*, which appears in the Bronze of Botorrita. The ancient Basque word for 'silver' could have been **urrezuri* (literally 'white gold'), which still survives in the Biscayan dialect, in which 'gold' is called *urregorri* (*gorri* 'red').

zuhain 'hay, fodder, tree'. Related to *zur* 'wood'. (Michelena, 1950b:200).

zur 'wood'. The final *-r* is secondary, determined by its final position. The root **zun* could explain *zunai* (*zuhain* 'hay, fodder, tree') and *zunar* (*zumar* 'elm') (Michelena, 1950b:199-200).

Comments on "Evaluation and Conclusions"

Trask is convinced that any prefix for the names of body parts never existed. Nevertheless, as he says, even Michelena admitted the possibility that the *b-* of most body parts might derive from *bere* 'his/her own'. According to Txillardegi (1980:28), this *b-* is a fossilization of the number *biga* or *bi* 'two', as most of these parts are duplicated. (This excludes from the list, of course, *buru* 'head', *bizkar* 'back' and others, as even the Basques only have one head and one back). Therefore, this prefix might be a remainder of a dual in Basque.

I would like to introduce another "nuance" in one of Trask's sentences, when he says that "Basque is, apart from a few prefixes involved in constructing verb forms and two or three prefixes acquired recently under Romance influence, exclusively suffixing." This is true with regard to old Basque, but today's Basque, like all modern languages, has developed a lot of prefixes, as it is the only way of becoming a useful language in all respects (Mujika, 1978).

Conclusions

After having read Trask's article, one can wonder why so many important authors have backed the Basco-Caucasian theory, in its different manifestations. It is clear that "common" vocabulary is limited and that the grammatical elements, such as declination, do not have too much to do with each other. Even the number of consonants is completely different (16 at most in Proto-Basque, 180 at least in Proto-North Caucasian, as Trask himself reminds us). In my opinion, there are only two linguistic reasons: ergativity and the verbal agreement, which is not enough to support this theory. As Trask explains, following Michelena, the linguists who back the Basque-Caucasian theory (and its variants) are "simply *assuming* that Basque and Caucasian must be related". Why?

In my view, because there are extra-linguistic reasons. I would dare to say even some mythical reasons. Since at least the XVIth century, Basque intellectuals have regarded Basconia as the new Eden and their language as the first one. Some "coincidences" (not fewer than those which are mentioned by several supporters of the Basque-Caucasian theory) helped them to get convinced. I have already mentioned some: the root *-ba* of the kinship terms and the false suffix *-eka* or *-ika* were compared to Hebrew words. But they were not the only ones: the name of Mount Ararat (where Noah's ark settled), might be the same as Aralar, one of the most important mountains of Basconia, which even today maintains a sacred meaning. Or the two Iberias, one the modern Spain and Portugal, the other one in the Caucasus. Basque authors managed to "find" hundreds of these cognates. I wonder whether the

Basco-Caucasian theory is not merely an attempted "scientific" elaboration of the ideas of the Basque intellectuals of the Renaissance, whose influence, through Humboldt and Romanticism, spread throughout the whole world.

In my opinion, at the level of our present-day knowledge (which, obviously, does not mean that this situation cannot change in the future), including Basque in Dene-Caucasian involves the same effort as Prokroustes, the famous Greek thief, stretching or shortening his victims to make them fit into his bed. In the same way, the roots of Basque are cut or lengthened to let them fit in a macro-family in which they are forced to be included without another option.

In fact, although some scholars have forgotten it, Basque is a hybrid language: like in English, more than half of its vocabulary (some authors raise it to three quarters) derive from Latin and Romance. Euskara has shown (Mujika, 1982) a great capacity for adaptation to a hostile environment. If the Basque language had refused to incorporate new words from other languages, we would be speaking about it as we do about Sumerian, Iberian, or Tocharian: as we do about a dead language. Because, as Michelena used to say (1954b:221), the true mystery about the Basque language is not that of its origin, but that of its conservation.

References

Corominas, Joan. 1972. *Tópica hespérica. Estudios sobre los antiguos dialectos, el substrato y la toponimia romances*, Gredos, Madrid, 2 volumes.

Kintana, Xabier. 1981. "Euskal eta Kartveliar izen batzuren kidetasunaz", in *Euskalarien nazioarteko jardunaldiak*, Euskaltzaindia, Bilbo, pp. 261-267

Michelena, Luis. 1949. "Voces vascas", in *Sobre historia de la lengua vasca*, Anejos del Seminario de Filología Vasca "Julio de Urquijo", San Sebastián, 1988, vol. I, pp. 448-457.

———— 1950a. "Notas etimológicas vascas", in *Sobre historia de la lengua vasca*, Anejos del Seminario de Filología Vasca "Julio de Urquijo", San Sebastián, 1988, vol. I, pp. 458-466.

———— 1950b. "De fonética vasca. La aspiración intervocálica", in *Sobre historia de la lengua vasca*, Anejos del Seminario de Filología Vasca "Julio de Urquijo", San Sebastián, 1988, vol. I, pp. 190-202.

———— 1954a. "De onomástica aquitana", in *Lengua e historia*, Paraninfo, Madrid, 1985, pp. 409-445.

———— 1954b. *La lengua vasca.* Biblioteca Vascongada de los Amigos del País, San Sebastián (2nd edition).

———— 1956. "Guipuzcoa en la época romana", in *Sobre historia de la lengua vasca*, Anejos del Seminario de Filología Vasca "Julio de Urquijo", San Sebastián, 1988, vol. I, pp. 138-155.

———— 1964. "Sobre el pasado de la lengua vasca", in *Sobre historia de la lengua vasca*, Anejos del Seminario de Filología Vasca "Julio de Urquijo", San Sebastián, 1988, vol. I, pp. 1-73.

———— 1970. "Nombre y verbo en la etimología vasca", in *Palabras y textos*, Universidad del País Vasco, pp. 283-309.

———— 1972. "Etimología y transformación", in *Lengua e historia*, Paraninfo, Madrid, 1985, pp. 296-308.

———— 1976. *Fonética histórica vasca*, Seminario Filología Vasca "Julio de Urquijo", San Sebastián (2nd edition).

———— 1979. "La langue ibère", in *Lengua e historia*, Paraninfo, Madrid, 1985, pp. 341-356.

———— 1987- . *Diccionario general vasco*, Real Academia de la Lengua Vasca, Bilbao (7 vols.).

Mujika, Luis María. 1978. *La prefijación, clave del euskera técnico y urbano*, Ediciones Vascas, San Sebastián.

———— 1982. *Latina eta erromanikoaren eragina euskaran. Euskal lexikoaren azterketa bideetan*, Sendoa, Donostia.

Perurena, Patziku. 1992. *Koloreak euskal usarioan.* Erein, Donostia.

———— 1993. *Euskarak sorgindutako numeroak.* Kutxa Fundazioa, Donostia.

Ruhlen, Merritt. 1987. *A Guide to the World's Languages. Volume 1: Classification.* Edward Arnold, London-Melbourne-Auckland

Sarasola, Ibon. 1984-1995. *Hauta-lanerako euskal hiztegia.* Gipuzkoako Aurrezki Kutxa Probintziala, Donostia (9 vols.).

Tovar, Antonio. 1959. *El euskera y sus parientes.* Minotauro, Madrid.

Txillardegi. 1980. "La lengua vasca", in *Historia de Euskal Herria*, Ediciones Vascas, San Sebastián, vol. IV, pp. 1-43.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

RESPONSE TO THE COMMENTS

R. L. Trask
School of Cognitive and Computing Sciences
University of Sussex, Brighton BN1 9QH, England
larryt@cogs.susx.ac.uk

This is a response to the comments of Zabaltza, Tiffou, de Grolier, Wescott, Hualde, Pulleyblank, Ruhlen, and Bengtson. The editors and I are hopeful that further comments will be received; if they arrive in time, I'll deal with them in a postscript. The commentators have made quite a few points, both general and specific. Naturally, for lack of space, I can't respond to every single one of these points, and so I shall confine myself here to addressing the general points plus a few specifics of particular interest.

1. Initial remarks. There was a slight misunderstanding between the editors and me. When I sent in my paper, I had thought I was sending in only a preliminary draft which I would have a chance to correct before the corrected version was sent out to the readers. As it happens, that first draft was sent out, and all readers have commented upon that. I will therefore begin by rectifying a few errors and omissions in that draft, including those now brought to my attention by the readers.

Section 1. Hualde points out that there are scholars who would deny my assertion that Iberian cannot be an ancestral form of Basque or even a close relative of it. True, such scholars exist, and some of them are eminent: the distinguished Catalan linguist Joan Coromines, for example, has recently been assuming that Aquitanian (pre-Basque) and Iberian were more or less indistinguishable, and has been invoking Basque etymologies with some freedom in studying the place names of the former Iberian-speaking region. The arguments advanced by such specialists, however, while seemingly plausible, are noticeably *a priori:* they represent a scenario which *might* have happened, but which is so far unsupported by any significant evidence. Indeed, their very success in "identifying" certain Iberian morphs with morphs in modern Basque seems to me to be a potentially fatal weakness: we are apparently required to believe that certain elements of Iberian have survived down to the present day in Basque absolutely unchanged, while everything else has been transformed utterly beyond recognition. I do not find this easy to accept.

I failed to mention the recent work on Basque and Kartvelian by Jan Braun (1981, 1985, 1994) and Xabier Kintana (1981). Braun expressly defends a remote genetic link, while Kintana is far more cautious in interpreting his comparisons. In my view, this work, like the Basque-Berber work of Schuchardt and of Mukarovsky, and indeed like all work involving Basque, suffers from the same shortcomings as the work discussed here.

As Blažek points out, there are some additional proposals in the literature for relating Basque to something else. Some of these were genuinely unknown to me, and I thank him for bringing them to my attention. Others I was aware of, but I haven't been able to get my hands on them. In any case, my list was never intended to be exhaustive, but only representative.

Concerning my remarks about the dissimilar morpheme structures of Basque and North Caucasian, I stand corrected (by Blažek and Bengtson) on the pervasiveness of monoconsonantal roots in North Caucasian; my declaration was rash, and I'm happy to withdraw it. I guess it's not wise to make assertions about languages you're not well-acquainted with. Note, however, that the minority of Basque morphemes containing only a single consonant are overwhelmingly disyllabic. Disyllables are the norm in Basque, and monosyllables are not at all common (though see now Lakarra 1995 for an argument that *all* Basque morphemes might have been monosyllabic at some very remote stage of the language).

In connection with my observation that a Basque-North-Caucasian link would entail a catastrophic collapse of the consonant system in Basque, Blažek points to Dyen's reconstruction of 46 consonants for Proto-Austronesian, which has been reduced to as few as eight consonants in some Polynesian languages. But Dyen's rich reconstruction is by no means generally accepted. Blust (1991) posits only 24 consonants for Proto-Polynesian and declares that few if any would now recognize more. Moreover, on examining the relevant papers in the journal *Oceanic*

Linguistics for the last ten years or so, I find that every single author posits no more than 23 or 24 consonants, and in one case as few as eighteen. Admittedly, 24 consonants to eight is still a startling degree of attrition, but hardly comparable with a hundred-odd to sixteen. And recall that pre-Basque had sixteen consonants only in intervocalic position; in all other positions, the number of contrasting consonants was very much smaller, as few as six word-initially and probably only five word-finally.

Bengtson now informs me that Starostin has proposed a very much smaller set of consonants for Proto-North-Caucasian. This is interesting news, and I can only wish that Starostin's work were published and hence generally available. I note, however, that, if Caucasic can be shown to have had only a modest number of consonants, then systematic correspondences between Caucasic and Basque should be very much easier to identify, if the comparison is valid.

Section 2. Hualde queries *antzara* 'goose'. In fact, *antzar is* the western form, *antzara* the central form, while eastern dialects have *antzare* or *antzera.* Hualde also adds L *arno* 'wine' under item [5] in this section.

As Hualde points out, the list of Romance varieties which have contributed words to Basque is garbled: it should read Castilian, Aragonese, Navarrese and Occitan (especially Gascon); the published version partly corrects this.

For Hualde's remarks on initial *k-*, see section 3 below.

Hualde is quite correct to point out that the aspiration cannot follow a word-initial liquid or nasal; I carelessly omitted this information in my summary.

2. Comments on particular comparisons. Just as the readers lacked the space to comment in detail on all of the 317 comparisons, I lack the space to respond to every one of their comments. I will therefore confine my attention to the ones I consider the most interesting and to the ones that might strengthen the comparativists' case.

Item [2]. Hualde disposes of Bengtson's objections to my account of Basque *abere* 'animal', but see further below.

Item [3]. The original mistake here is Čirikba's in citing the non-existent **abets* for 'voice', but I compounded it by assuming, without checking, that this was yet another of Sabino Arana's eccentric neologisms parallel to his notorious "*abesti*" 'song'. I am happy to confirm that, as pointed out by Zabaltza, the word does not exist at all. Zabaltza adds that the genuine word *ahots* was coined by Larramendi in the eighteenth century; this may be so, but it is not confirmed by Agud and Tovar (1987-).

Item [4]. Hualde points out that my observations about Basque *adar* 'horn' do not prevent it from being of "Macro-Caucasian" origin. Agreed.

Item [5]. Basque *adin* 'age', 'judgment', and other senses has long been a puzzle. In my paper I draw attention to one possible etymology, but Zabaltza has reminded me of another suggested by Michelena, involving the distal demonstrative stem and the adjective-forming suffix *-din.* I confess I prefer this one, though I am slightly troubled by the Aquitanian form cited by Blažek. On balance, though, I find it hard to concur with Blažek's judgment that this comparison "belongs to the most convincing".

Item [8]. Bengtson objects to my citing (or rather Tovar's citing) Latin AVULE as the source of *agure* 'old man', on the ground that the Latin word is explicitly a vocative. But the whole point is that old men are frequently *addressed* as 'grandpa', and I don't see the problem.

Item [11]. On the status of Basque [h], see section 3 below.

Item [23]. It occurs to me that Čirikba's alleged and unexemplified Basque "plural" suffix **-ar* which I here reject as non-existent may in fact have been intended as a reference to the noun-forming suffix *-ar* found in a few words:

galtza(k) 'trousers' (and other garments), *galtzar* 'side of the body' (and other senses). But this suffix is in no sense a "plural" suffix.

Item [27]. As Hualde points out, the accepted reconstruction of Basque *arrai(n)* 'fish' is **arrani*, which improves the putative match with Coptic on the phonological side, though it doesn't do much for the semantics. Joseba Lakarra has suggested (p.c.) that the Basque word might be an ancient participle; cf. Spanish *pescado* 'fish', derived from a participle.

Item [32]. My citation of **"stager"** for Batsbi **stagar** 'maple' was a typo, picked up by Bengtson.

Item [39]. Bengtson rejects Michelena's (and Luchaire's and Meyer-Lübke's) etymology for *azeri* 'fox' as "absurd and fantastical", in spite of the abundant evidence for it sketched out in my paper. He gives two reasons: (1) the Z form *azeri ~ axeri* has no nasalized vowel; (2) the B forms *azegari ~ azagari,* with a /g/ that I consider secondary, must be the most conservative forms because B is a peripheral dialect, and peripheral dialects never innovate. First, Z does not *invariably* retain the earlier nasalized vowels; it merely *usually* does so. Consider *ohol* 'plank', derived by Michelena from **onol*. R *õl* indeed shows the expected nasal vowel, but Z doesn't: it has *ohol,* with no nasal vowels. Second, if, as Bengtson maintains, peripheral dialects like B are prohibited by law from introducing any innovations at all, and prohibited in particular from inserting a /g/ to break up vowels in hiatus, then how does he explain cases like B *agate* 'duck', which is *a(h)ate* in most other dialects (with nasal vowels in the east) and which is borrowed from Latin ANATE? Further, to the evidence cited in my paper, I should add that **azenari* is rather directly recorded in the name of the Aragonese town *Acenarbe* and less directly recorded in the name of the Bizkaian town *Zeanuri* (< **aze(n)ar-uri*; *uri* 'town') (Michelena 1973).

Item [47]. The two Basque words *barret* and *barrasaket,* both 'sterile', are almost certainly of Occitan origin, as pointed out by Zabaltza and Hualde; this may possibly mean that they should be removed from the discussion here, but I'm not sure.

Item [49]. I accept Hualde's comments about the vegetation.

Item [55]. Hualde's observation about the frequency of Basque /m/ in "expressive" reduplications is important; compare item [215], which I dismiss for the same reason.

Item [59]. I neglected to add that the existence of an ancient morpheme **bel* 'dark, black' has very recently become generally accepted among vasconists; to Zabaltza's citation (under item [112] *gibel* 'liver') of the work of Perurena, I can add Lakarra (1995). This is enough to destroy the proposed comparison.

Item [60]. I cannot accept Hualde's suggestion that Basque *ber-* 'same, self' constitutes a reasonable match for words meaning 'thresh' and 'split'.

Item [65]. Zabaltza points out that several vasconists have seen *bete* 'full' as originating in *bat* 'one' (< **bade* or **bada*). If this etymology is correct, it would destroy the comparison.

Item [69]. Basque *bihar* of course means 'tomorrow' and not 'yesterday', as pointed out by Zabaltza; this was a typo, and it has been corrected (I hope) in the published version.

Item [80]. Ruhlen persists in maintaining that Basque *d-* is a "third-person marker". But it is not: it is a present-tense marker. This view, advanced only cautiously in my paper, is now in fact probably universally accepted among vasconists: see for example Gómez and Sainz (1995). De Rijk (1995) has even proposed a source for this prefix within Basque. And I have recently discovered that the same insight was put forward independently by Oregi Aranburu (1974).

Item [98]. Basque *esne* 'milk', with its unique cluster, has an attested variant *esene,* and in all likelihood, as pointed out by Hualde and Zabaltza, it derives from a longer form, possibly **esende,* as suggested by Michelena. This conclusion possibly improves the matches with the other languages, though hardly dramatically.

Item [104]. Though the comparison is already hardly viable, I should point out further that there are excellent grounds for regarding Basque *ezti* 'honey' as an ancient compound. Not only is *-ti* a very familiar noun-forming suffix, but the element *ez-* appears to recur in *ezko* 'wax', and perhaps also in *erle* 'bee' (< **ez-le* ?, perhaps 'honey-maker', where *-le* is the ancient agent suffix), as suggested by Michelena (1977:366) (and also, as Zabaltza points out, by Kintana).

Item [112]. Zabaltza is quite right to point out that *gibel* 'liver' is a compound involving **bel* 'dark'; the first element is possibly shared with *giharre* 'lean meat', *gizen* 'fat, fatty meat' and *giberri* 'head of cattle' and may perhaps have meant 'meat' (Lakarra 1995:192). This compound status is enough to destroy the proposed comparison.

Item [114]. A clarification. There is abundant evidence that Basque *urdin* formerly covered all of green, blue, and gray. For example, a certain mushroom with a bright green underside is called *gibelurdin* in Basque (*gibel* 'back'). It appears that the introduction of Romance loans like *berde* 'green' has narrowed the meaning of the Basque word. Exactly the same thing has happened to Welsh *glas*, formerly 'green, blue, gray', but now restricted to 'blue' in all but a few set expressions.

Item [125]. I neglected to add that Basque *inguru* 'vicinity' has a less common variant *ingiru*, which more directly reflects its origin in Latin IN GYRU 'in a circle'. Bengtson queries my etymology, asking "how did it get from here to there?" Well, the word is abundantly attested in Latin and Romance, even in modern Italian, as pointed out by Hualde, and the Basque Country was, after all, part of the Roman Empire from the first century BC onwards. What is the problem?

Item [129]. José Ignacio Hualde has recently suggested to me (p.c.) that Basque *handi* 'big' might be an ancient compound involving the venerable and well-attested adjective-forming suffix *-ti.* I find this idea exceedingly plausible, especially since there are several other adjectives and quantifiers which look as if they are formed in the same way, such as *guti* 'not much, not many' and *guzti* 'all'. If this proposal is correct, the comparison is destroyed.

Item [135]. See section 3 below for some discussion of the initial [h] in these demonstratives, including a proposal which would destroy the proposed matchups.

Item [151]. Bengtson rejects the derivation of Basque *horma* 'ice; wall' from Latin FORMA, and prefers his Burushaski match **Ramu.** Well, whatever one might think of the Latin etymology, the fact is that the word is firmly attested as *borma* in both senses in Oihenart, which is enough to ruin the comparison. As for the semantics, the descendants of Latin FORMA exhibit an extraordinary range of senses in modern Romance, everything from 'shoemaker's last' to 'ditch'; many of the senses have to do with molding, shaping, or outlining something. Its Spanish descendant *horma* means 'drystone wall', which accounts easily for one of the Basque senses. In its earliest attestation in the 'ice' sense, Basque *borma* means, not 'ice' in general, but specifically a thin layer of ice which coats something like a mold — and I don't see any great problem.

Item [156]. I accept Hualde's observation about the different nature of the terrain north of the Pyrenees, but I don't think it adversely affects what I say here: quite the contrary, in fact.

Item [161]. I accept Zabaltza's correction, and withdraw my suggestion about *hamaika.* I didn't really think I was going to get away with that one.

Item [194]. Hualde and Zabaltza are right to point out that there is a long-standing debate over the southern limits of Aquitanian speech in Roman times. Here I adopt a conservative view, since there is little or no linguistic evidence for the presence of Aquitanian-speakers as far south as the Ebro Valley, and most specialists posit Celtiberian or even Iberian speech for this area. The place-names in *Il-* cannot be taken as sure evidence for Aquitanian speech; place-names in *Ili-* ~ *Ile-* ~ *Ilu-* are numerous and are found throughout the Iberian-speaking area in eastern and southern Spain, as well as in Aquitania (see the map on page 89 of Anderson (1988), compiled by Untermann).

As Hualde points out, I was too rash in asserting that initial *k-* is unknown in Aquitanian; on this, see further in section 3 below. But the idea that Aquitanian might have had a word **kala,* meaning 'castle' or even 'house', is sheer fantasy. If **kala* had existed in Aquitanian at all, it would have come into Basque as **gara* — and naturally I think at once of Basque *gara* 'high place, elevation', well attested in toponyms. It would be interesting to know if Calahorra is located upon a height. So far as I can tell from the best map I have, it is not: it is instead located at the confluence of two rivers, in terrain that looks unrelievedly flat. But I've never been there. Nor would such an etymology be of the slightest assistance to Bengtson in any case.

Item [197]. See now my remarks on Basque *karats* 'bitter' in section 3 below

Item [199]. In response to my dismissal of this comparison, Ruhlen somewhat surprisingly comments that my "only objection" is that Basque *korotz* ~ *gorotz* 'dung' is a loan from Romance. I would have thought that this one objection was quite sufficient. Since Ruhlen (and Bengtson) also complain later about the "bizarre semantic shifts" I appeal to, perhaps I'd better discuss this word in detail.

Corominas proposes that the Basque word derives from a Romance development of Latin CROCEA 'saffron-colored'. This word in attested in post-classical times as meaning 'saffron-colored garment' (it's in the Vulgate). Its Romance descendants in northern Spain, including Old Spanish *croça*, are abundantly attested as denoting various rude garments of a yellowish color, always in a rustic context; the word was applied to caps, hoods, and cloaks, and was possibly first applied to a broad-brimmed straw hat worn as protection against the rain. Significantly, the word was also applied to a covering placed over a haystack to keep it dry — and such coverings were made of dung, mixed with straw. Corominas therefore proposes that Basque-speakers borrowed this (attested) word in this (attested) sense, and merely generalized its meaning from 'dung used to cover a haystack' to 'dung' (in general) — not a very "bizarre" semantic shift, after all.

As Corominas himself admits, this etymology is by no means certain, but it is very plausible, and it is backed up by a certain amount of evidence. So:

(1) The Basque word doesn't look like a native word; it looks like a loan word;

(2) It has a very plausible source in an attested Romance word;

(3) It has *exactly* the phonological shape it would have if it were borrowed from that Romance source;

(4) We already know that Basque has borrowed thousands of words from Romance, and so finding one more is hardly going to raise any eyebrows. On balance, then, which is more plausible: that the Basque word is indeed a loan word, or that it continues a 7000-year-old Dene-Caucasian etymon?

Even if Corominas's etymology turns out to be wrong, the central point here is that it is reckless to hypothesize etymologies from thousands of miles away without first considering simpler explanations. Other sources are available closer to home, and these should be preferred to "audacious" hypotheses as a matter of principle and practice.

Item [200]. As Bengtson points out, Old Spanish *coma,* being rare and elevated, is a most unlikely source for Basque *kuma.* On reflection, I have realized that this strictly eastern word must derive instead from the synonymous Occitan *como* ~ *coumo* ~ *couma*, which is abundantly attested in that language (Mistral 1968; Meyer-Lübke 1935) and which is, of course, cognate with the Old Spanish word. As for the (localized) G form *kima*, this is of no

consequence, since the sporadic fronting of /u/ to /i/ is common in all varieties. For example, *umore* 'humor', which surely even Bengtson will admit as a loan from Latin HUMORE, appears as *imore* in some varieties of HN and LN (central dialects, as it happens), and there are dozens of other such cases (Michelena 1977: ch. 3).

Item [210]. I confess I was unaware of the existence of Latin MARRUBIU 'horehound', pointed out by Zabaltza, and I now find that precisely this etymon for the Basque word is cited as early as Meyer-Lübke (1935). Oh, well, you're never too old to be spectacularly ignorant. This source, of course, at once destroys the proposed comparison.

Item [214]. I believe I have identified the source of the confusion over the alleged (but non-existent) Basque **maño* 'masculine'. The Basque word for 'mule' is *mando,* generally thought to be a loan from Celtic (the word is attested throughout western Europe and the northern Mediterranean coast; it is applied to various animals, especially horses, and the meaning 'mule' appears to be specifically Celtic). If this word formed an expressive diminutive, this would be **mañddo*, which in western dialects could be reduced to **maño.* Neither of these forms is listed in any of my dictionaries, but I would not be surprised to find that they exist, since they would be unremarkable. Now the Basque word has another meaning: 'sterile'. One of the Spanish words for 'sterile' (and also 'mule') happens to be *macho* — and this, of course, is homophonous with the much better-known Spanish word meaning 'male, masculine, virile'. Hence I conclude that somebody has probably misunderstood a Spanish gloss of a diminutive of the Basque word for 'mule', 'sterile', and amusingly taken it as denoting almost the very opposite. (And I see now that Bengtson has independently reached the same conclusion.)

Item [220]. Bengtson denies my reconstruction of the Basque word for 'pith' as **bune* ~ **(h)une* < **fune,* on the *a priori* ground that peripheral forms are *always* more conservative than central forms. I don't think I need to waste any time discussing this absurd dogma. I will content myself with noting that, in my original paper, I overlooked one more modern variant of the word: *fuin* (or *fuiñ*) in Lapurdian. This would appear to make the case for original **fune* overpowering, since it alone can account straightforwardly for *all* the attested variants. And I would further add that the word is also attested in the sense of 'spinal cord', which perhaps eases the semantic awkwardness of invoking Latin FUNE 'rope'.

Item [230]. See section 6 below for my response to Ruhlen's comments on Basque *odol* 'blood'.

Item [250]. Eastern allative forms like R *etxeara* are almost certainly late innovations, obtained by attaching the new allative ending *-ra* to an NP bearing the determiner.

Item [251]. The analysis of Basque *sabel* 'stomach' as containing **bel* 'dark', as suggested by Zabaltza, is more than plausible, though not certainly correct. If it is correct, the comparison is destroyed.

Item [254]. I am unable to accept Ruhlen's speculation that Basque *sahats* (and variants) 'willow' might derive from **sakats*. There is no parallel for the loss of intervocalic /k/ in Basque; compare the native words *akats* 'scratch, nick, defect' and *sakon* 'deep', which have no variants without /k/. If there are difficulties with the obvious **sanats*, then a good guess would be **sarats* or **sabats*, either of which could easily yield the attested variants. Moreover, both an element **sar-* and an element **sab-* are well attested in botanical terms: *sarale* ~ *saraille* 'hay', *sarama* 'brushwood', *sarasabi* 'nursery for plants or trees', the hapax *sarbagorri* 'beet' (*gorri* 'red'; the element *sarba-* is otherwise unattested), *sarbegi* 'small branches, twigs, leaves on a branch' (*begi* 'eye'), *sarga* 'branch; esparto grass', *sarobe* 'type of chestnut', the hapax *saroi* 'forest', *sarbi* ~ *sabi* ~ *sabitegi* 'nursery' (*-tegi* 'place'), *sabi* 'network of fine roots on a tree', G *sabiko* L LN *sabuka* Z *sabukitze* 'elder tree' (Z *-(a)tze* 'tree'). (Also possibly relevant is *sasi* 'bramble'; this is attested in one early text as *zarzi,* but Michelena (1977:284) considers that **sarzi* might be a plausible reconstruction, with the usual Basque sibilant harmony. L *saradon* 'variety of laurel' and the hapax *sarandoi* 'celandine' are loans from Gascon and can hardly be relevant here.) These words have no etymology, but it is difficult not to recognize an element **sar(a)* as being present in some of them and perhaps also an element **sa(r)bi* (a derivative of **sar(a)*?). Of course, there is nothing certain here, but a link with some of these

other indigenous words seems vastly more plausible than the fanciful phonological gyrations proposed by Ruhlen expressly to try to tie up Basque *sahats* and Burushaski *šAsk*.

Item [255]. On reflection, I believe that Zabaltza is almost certainly correct in seeing Basque *samin* 'bitter' as containing the common element *min*, which is 'pain' in isolation but which in compounds frequently means 'spicy, hot'. If correct, this analysis would destroy yet another comparison.

Item [267]. Hualde queries my dismissal of Basque *-ta* as a locative case-marker. A full discussion of my views here will have to await the publication of Trask (forthcoming c). Briefly, though, I suggest the following. The Basque morph *-eta* may be of Latin origin, but need not be. It was probably a collective suffix in origin, but it naturally came to be reinterpreted as a plural. The sequence *-eta-n*, plural plus locative *-n*, was then analyzed as containing the ordinary oblique plural marker *-e-*, and this vowel was therefore removed to construct an innovating non-plural locative ending *-tan*. A key point here is that the Basque local cases are complex and anomalous in their formation, and there is every reason to believe that they have been stitched together in comparatively recent times from bits and pieces of various origins. Recall too that the modern singular and plural definite inflectional paradigms are of very recent origin in Basque; they could not have existed before the creation of the definite article, probably in post-Roman times.

Item [269]. See section 5 below for my response to de Grolier's remarks about Basque *tu* 'spit'.

Item [277]. I omitted to include the following further information. The word for 'nine' is *bede(r)atzü* in the Zuberoan dialect and *bedratzu* in Roncalese, both plainly from **bederatzu*. Since there is no parallel for the unconditioned development of /i/ to /u/ in these or any other dialects, it follows that **bederatzu* must be the original form of the numeral, and that the *bederatzi* of the central and western dialects must reflect contamination by the preceding *zortzi* 'eight' (Michelena 1954:277). (Such contamination between consecutive numerals is, of course, exceedingly well attested in Indo-European languages.) Quite apart from the points made in my paper, this fact is itself enough to destroy the proposed comparison.

Items [282, 283]. Bengtson rejects Michelena's analysis of *ukarai* 'wrist' and *ukondo* ~ *ukalondo* 'elbow' as compounds involving the archaic *uko* 'forearm', the first with *garai* 'high part' and the second with *ondo* 'bottom', a loan from Latin FUNDU. His objections are these: (1) Basque *uko* is "a hypothetical Basque word for 'hand, arm' [*sic*]; (2) no other European language has a word for 'elbow' that is a half-native, half-foreign, compound of this type. First, *uko* is not hypothetical: it was used by the 17th-century writer Oihenart, a very reliable source who did not coin neologisms, and it can be found in Azkue's dictionary. Second, I cannot see any force in the failure of other European languages to follow Basque in forming their words for 'elbow'. Basque has lots of body-part names constructed with spatial nouns, including the loan *ondo*: *hankagain* 'hip' (*hanka* 'leg' + *gain* 'top'), *izterrondo* 'groin' (*izter* 'thigh' + *ondo*), *sudurgain* 'bridge of the nose' (*sudur* 'nose' + *gain*), *galtzarpe* 'armpit' (*galtzar* 'side of the body' + *-pe* 'below'), *eskubarne* 'palm' (*esku* 'hand' + *barne* 'interior'), and others — and surely the facts of Basque are more important here than the facts of Italian or Polish. Moreover, English has lots of body-part names (though not 'elbow') compounded from both native and borrowed elements: *kneecap*, *armpit*, *earlobe*, *index finger*, *rib cage*, *soft palate*, *hairline*, *belly button*, *canine tooth*, and perhaps *eyelash*.

Item [305]. Bengtson rejects the reconstruction of *ziho* 'oil, fat' as **zino* because Z has no nasal vowels in this word. True, but R, the other dialect which usually retains nasal vowels, *does* have a nasal vowel. It's a good idea to consider *all* the evidence.

Item [308]. As Hualde and Zabaltza point out, there is Celtiberian evidence for something close to Basque *zilar* 'silver', making the possibility of a loan from Berber even more remote.

In his reply, Bengtson adduces a number of new comparisons involving Basque and the other languages. I cannot presume upon the good will of the editors to the point of discussing every one of these in detail, and so I shall

content myself with a brief demonstration that these new comparisons are just as deeply flawed as the earlier ones. The new items are not numbered, and so I give them here in alphabetical order; I also include one new comparison advanced by Ruhlen.

Bq **abere** 'animal': Bengtson rejects the universally accepted derivation of this from Latin HABERE 'have', on the grounds that Basque does not borrow Latin infinitives, that the word means only 'animal' in general, and not 'domesticated animal' or 'possessions', and that its combining form is *abel-*. But Basque didn't have to borrow the infinitive *qua* infinitive: it was already in widespread use in early Romance as a noun. The descendants of the Latin word mean 'animals for agricultural work' in Murcia, 'cattle' in Segovia, 'head of cattle' in Catalan, 'bovine animal' in Galician, 'animals, flock' in Occitan, 'animals' in Normandy, and so on — so there is no morphological or semantic problem. Moreover, the Basque word means 'animal' (in general) only in some places; elsewhere it means specifically 'domesticated animal', 'animal involved in husbandry', 'cattle', 'bovine' or even 'equine'. Finally, the combining form results from rule inversion: compare eastern Basque *zamari* 'horse', a loan from Latin SAGMARIU 'pack-horse', whose combining form is *zamal-*, as in *zamaldun* 'horseman'.

Bq **ahari** 'ram' (not *'sheep'): This word has nasal vowels in the east, and so we can securely reconstruct **anari*, destroying the comparison.

Bq **alme** 'side': The meaning 'side' is marginal; most usually, this word denotes a cut of meat, more or less what we call 'flank steak'. It is one of a large group of words in *al-*, all meaning 'side' or something closely related. The second element is almost certainly the common word *mehe* 'slender, thin' (< **bene*). The related *almehaka* 'flank, side' appears in many western varieties as *meaka* or *beaka*, also a derivative of *mehe*. Basque *alme* is thus a compound, and its /m/, as always, derives from */b/, destroying the comparison. On all this, see Agud and Tovar (1988-) under *almaka* and *albo*.

Bq **borroka** 'fight': The various attempts at giving this word a Romance etymology have been unconvincing. The final *-ka* is surely the suffix which occurs so frequently in the names of activities, games and contests; this is enough to ruin the comparison. A plausible source is *borro* 'ram bred for fighting'; the transparent **borro-ka* 'ram-fight' might easily have been extended to 'fight' in general.

Bq **ehun** 'linen': This widespread and clearly ancient word actually means 'cloth' in most places, rather than specifically 'linen', which possibly improves the comparison (though I strongly suspect that 'linen' is in fact the earlier meaning). However, a difficulty is that this noun is derived from the participle of the verb *ehun* 'weave', and hence it means literally 'woven (stuff)'. In all likelihood, the original form was **enun*, but the eastern dialects do not distinguish nasal from non-nasal vowels before a following /n/.

Bq **erdera** 'foreign' [*sic*]: The Basque word *erdara* ~ *erdera* means only 'foreign language', and it is transparently bimorphemic: it contains the suffix *-ara* ~ *-era* which also occurs in *euskara* ~ *euskera* 'Basque language'. This destroys the match.

Bq **-ez-** (adjective) [*sic*]: Ruhlen extracts this from *arrezko* 'male', a derivative of *ar* 'male', and compares it with adjective-forming affixes in other languages. But this analysis is hopelessly wrong. The exceedingly frequent Basque suffix *-ko* can be added to virtually any adverbial at all to derive an adjective-like modifier. One of the things it can be added to is an NP bearing the instrumental case-suffix *-z*: *egia* 'truth', *egiaz* 'in truth, truly', *egiazko* 'true'. The resulting compound suffix *-zko* has been generalized to derive from nouns adjectivals meaning 'having the character of', as in *urre* 'gold', *urrezko* 'made of gold, golden'. This is the source of *arrezko*, in which the /e/ is automatically inserted, as always in Basque, to break up the resulting consonant cluster. Ruhlen's **-ez-* does not exist.

Bq **habe** 'pillar, beam': This word has a simply extraordinary number of senses, most of them pertaining to 'support' or sometimes 'protection', in both physical and metaphorical senses. The physical senses, which are surely earlier, almost exclusively involve supports made of wood, and oddly enough the earliest attested sense of the word is simply 'tree'. Agud and Tovar take this seriously, and suggest that 'tree' is the original sense, with the senses of 'beam, rafter, pillar' then being derived from the obvious uses of tree-trunks in construction, in a manner familiar from the Romance languages, and they further suggest that *abar* 'branches' is a derivative of *(h)abe* with the familiar noun-forming suffix *-ar*. It is not clear whether any of this is of any assistance to Bengtson.

Bq **hanka** 'leg': This obvious loan is dismissed by Bengtson as follows: "[It] has nothing to do with Latin or Romance: it is usually explained as a loanword from Germanic, but what is the mode of transmission?" Very well. A Germanic word **hanka* 'haunch' is found widely in the Germanic languages (*e.g.*, Dutch *hanke* 'haunch'). This word was borrowed from Frankish into Gallo-Romance, whence it spread throughout most of western Romance (Spanish *anca* 'haunch', French *hanche* 'haunch, hip' [Old French *hanche* is the source of English *haunch*], and so on). In the French Basque dialects, *hanka* generally means 'haunch, rump'; in Lapurdian it means 'leg'; south of the Pyrenees it is variously 'leg, foot, paw'. Is this good enough?

Bq **har** 'worm': This word has a nasal vowel in the east, and Michelena reconstructs **anar*, destroying the comparison.

Bq **hausin** 'nettle': This word has a large number of regional variants, of which *(h)ausin* is specifically Lapurdian, and it has attracted a good deal of discussion. Some have seen it as a derivative of B *asu* 'bramble', another prickly plant. I am inclined to wonder whether the L form has not been influenced by Occitan *ausina* 'ilex', yet another prickly plant.

Bq **hauspo** 'lung': This word means 'bellows' throughout the Basque Country, with 'lung' being attested only in B; this last sense is clearly secondary, and it should not be cited as the primary meaning. The word appears variously as *(h)auspo ~ (h)aspo ~ (h)ausko*; some have seen it as a derivative of *hauts* 'ash', but most vasconists would probably follow Agud and Tovar in deriving it from *hats* 'breath' plus the frequent noun-forming suffix *-ko*.

Bq **hümoi** 'womb': This form is specifically Z; elsewhere, the word is *umoi*, and it variously means 'womb' or 'cradle'. It is a transparent compound of *ume* 'child' and *ohe* 'bed'.

Bq **hurolde** 'flood': The severely localized form *hurolde* has apparently been chosen merely to get the initial /h/ to make the comparison look better. In fact, even in the aspirating dialects, the word is normally *uralde ~ u(h)alde ~ uraldi ~ ualdi* 'spate, flood'. This is transparently a compound of *ur* 'water', which is enough to destroy the comparison; the second element is most likely *aldi*, which is extremely frequent as a second element in compounds and means 'occasion on which, or period during which, something happens' (compare *udaldi* 'summertime', from *uda* 'summer').

Bq **idulki** 'block of wood': This word is attested in this sense nowhere but in the 18th-century Bizkaian writer Añibarro. It is clearly a derivative involving the suffix *-ki*, which forms concrete nouns, though the first element is obscure: neither B *idun* 'wooden pole' nor B *iduri* 'powdered charcoal' seems obviously right, though either might be defended.

Bq **ilaski** 'moon': The nearly universal Basque word for 'moon' is *ilargi*, a compound of **iLV-* 'moon' and *argi* 'light'. The isolated *ilaski* (which is Salazarese, not "Lapurdian") should not be extracted and used exclusively in comparisons merely because it is more convenient. Though also an obvious compound of **iLV-*, this form has generally been considered a puzzle, but I would suggest a simple etymology: it is

merely the ordinary *ilargi* plus the very common noun-forming suffix *-zki* (also found in *eguzki* 'sun'); the development **ilargi-zki* > **ilart-zki* > **ilarzki* > *ilaski* would be absolutely regular.

Bq **katu** 'cat': Bengtson denies that this is a loan from Latin CATTU, suggesting that the word might be pre-Roman in Basque. Well, the *animal* might have been pre-Roman, but this very un-Basque-looking word has *exactly* the shape it would have if borrowed from Latin.

Bq **lama** 'flame': For this word too Bengtson doubts that Latin FLAMMA is the source. Remarkable. Perhaps he also doubts that Basque *lore* 'flower' derives from Latin FLORE, or that Basque *luma* 'feather' derives from Latin PLUMA.

Bq **lapa** 'burdock': This is a straight loan from Latin LAPPA 'burdock'.

Bq **medar** 'narrow': This is a curious central variant of eastern and western *me(h)ar* — and we know what Bengtson thinks about central forms. What has happened to his conviction that peripheral forms are always more conservative? The word is derived from *mehe* 'thin, slim' (< **bene*), probably with the word-forming suffix *-ar*. Compare the synonymous *mehats* ~ *meatx,* with another word-forming suffix, *-ts*, and see **alme** above. All this destroys the comparison.

Bq **ondo** 'bottom, side' (and other senses): Bengtson rejects my assertion (under item [283]) that this is a loan from a Romance development of Latin FUNDU 'bottom', and proposes a Caucasian cognate. To this end, he adduces Bq **zuhañ-ondo** 'foot of a tree', **bazkal-ondo** 'after breakfast' [*sic*], properly 'after lunch', and **ondo-amen** [*sic* with the segmentation] 'consequence', and remarks "Latin cannot explain all this". Perhaps not, but nobody is suggesting that Latin should be invoked to explain formations in Basque. The first compound is literally 'tree-bottom', and I can see no difficulty. The second merely illustrates the very frequent sense of 'after' which *ondo* has acquired in Basque; I see no difficulty with a semantic shift from 'bottom, side' to 'after', and I fail to see how a Caucasian word meaning 'joint, bone' provides greater illumination. Finally, *ondo* also has a common transferred sense of 'consequence', a natural enough development from 'after'. What Bengtson is citing is *ondoramen*, which consists of *ondora*, the allative of *ondo*, plus the abstract-noun-forming suffix *-men*, a borrowing from Latin -MENTU; the Z form *ondoamen* merely illustrates the usual loss of intervocalic /r/ in that dialect.

Bq **orga** 'wagon': This word means 'wagon, cart, oxcart', and its eastern form *orgã*, with a nasal vowel, together with its combining form *organ-*, allows us to reconstruct **organa*, destroying the matchup. In all likelihood, the word is, *contra* Bengtson, derived from Latin ORGANA, plural of ORGANUM; it is noteworthy that Azkue records that the Basque word is normally used in the plural even when speaking of a single wagon. Why is a semantic shift to 'wagon' more surprising than the English senses of *organ*, such as '(bodily) organ' or '(official) periodical'?

Bq **orthuts** 'barefoot': This is a simple compound of *oin* 'foot' and *huts* 'bare'. The /n/ ~ /r/ alternation in this position is ancient and regular, though of unknown origin; a /t/ often appears between elements in word-formation and has a well-understood source (compare *sutargi* 'firelight', from *su* 'fire' and *argi* 'light'); the reduction of the diphthong before a cluster is regular.

Bq **oski** 'shoe': This word has a similar origin: it is *oin* 'foot' plus the concrete-noun-forming suffix *-zki*. The development of **oinzki* to **orzki* parallels that of **orthuts,** and the simplification of **orzki* to *oski* is entirely regular (compare *uso* 'pigeon', from *urzo*, preserved in the east).

As can be seen, the new comparisons are in no respect an improvement upon the earlier ones. Curiously, Ruhlen chooses to devote most of his reply to merely listing nearly 70 of the comparisons cited in my paper, with no new

discussion, and even including a number which I have already shown to be plainly wrong. I have no idea why he thinks this is a useful thing to do.

3. The Pre-Basque phonological system. In my paper, as in all my work, I accept as essentially correct the reconstruction of the phonology of pre-Basque (roughly 2000 years ago) by Luis Michelena, as laid out in the book cited here as Michelena (1977). There is nothing in the least surprising about this policy. Michelena is beyond dispute the greatest scholar who ever worked on Basque, and his knowledge of the language vastly exceeded that of anyone else who has ever lived. His reconstruction, which occupies a book of some 600 pages, is abundantly and meticulously documented; it takes into account *all* the evidence, from the forms of Basque names and words recorded in medieval documents to the forms of the numerous modern dialects. The system he reconstructs for pre-Basque, while very different from the systems of Indo-European languages, is symmetric, economical, and appealing; far more importantly, it *works*. For example, it accounts very successfully for the forms of loan words from Latin and early Romance, and the few and simple phonological changes he posits as having applied to this system account with equal success for the forms occurring both in medieval documents and in the several modern dialects. As I point out in my paper, this is one of the finest pieces of historical reconstruction one could ever hope to encounter, and Michelena's conclusions are accepted as valid by all specialists today. Nothing short of the most substantial and astonishing new evidence could possibly disturb Michelena's reconstruction.

I am therefore more than surprised to find that some of the readers would dismiss this reconstruction with no more than a wave of the hand. Ruhlen, for example, dismisses this abundantly and meticulously documented reconstruction of pre-Basque phonology as nothing more than "[Trask's] preconceptions" — even though I had nothing to do with that reconstruction, and even though it is accepted as correct by all vasconists without exception. Indeed, Ruhlen appears to hold the entire body of professional work on the prehistory of Basque in utter contempt: for him, it's just something to be waved away whenever it gets in the way of his "audacious" hypotheses.

Blažek similarly takes me to task for invoking Michelena's reconstruction as though it were beyond question. Well, as far as any specialist can see at present, it *is* beyond question. (In fact, Blažek, like Ruhlen, describes Michelena's massive reconstruction as "only...[Trask's] own postulates of Basque historical phonology".) Blažek further invokes the chronology of Basque phonological development proposed by Guiter. Now, to start with, Guiter has no track record as a vasconist; moreover, assuming he is the same Henri Guiter who some years ago attempted to interpret the entire corpus of Dark-Age Pictish inscriptions in Scotland, including those which are so badly weathered as to be quite illegible, as a collection of flowery laments written entirely in twentieth-century Basque (Guiter 1968), one might reasonably look askance at any of his declarations about Basque. Be that as it may, the fact is that Guiter's version of Basque historical phonology is less than four pages long and is supported by no evidence. For the later stages, he accepts Michelena's posited changes, but he goes so far as to put hard dates on every one of them, dates which are almost entirely fanciful and which in some cases are clearly wrong. For the early stages, he departs radically from Michelena's reconstruction, positing instead a very different system, again with not a shred of evidence. Why should anyone take seriously this miniature flight of unsubstantiated fancy, when Michelena's huge and sober tome is available? The point is not what kind of pre-Basque phonological system would be convenient for hopeful comparativists; the point is what kind of system is required by the facts of Basque. Anything else is a waste of time.

But it is Bengtson who has the harshest words for what he calls the "Michelena-Trask Pre-Basque phonology". He describes Michelena's reconstruction as "*hypotheses*" [his emphasis] and "speculations" which I expect people to accept "*on faith*" [his emphasis]. I think he might like to reconsider this rash language: a reconstruction backed up by 600 pages of meticulous documentation and accepted as valid by all specialists is not the sort of thing that can reasonably be called "speculation", and it certainly doesn't have to be accepted "on faith": anyone is free to scrutinize Michelena's evidence, which involves appeals to more than 4500 Basque words and morphemes, as well as to some 400 Latin and Romance words.

Bengtson goes on to raise four objections to Michelena's reconstruction.

(1) "It is based on faulty logic: one possible interpretation of the evidence is adopted, ignoring other possible interpretations." Well, given Bengtson's insistence on clinging to his Dene-Caucasian etymologies for Basque words in the face of the plainest evidence to the contrary, I cannot see that he is in a position to advance this

particular criticism of the work of vasconists. In any case, the reason Michelena adopts his particular interpretation is that it is supported by the facts of Basque, while the "other possible interpretations" are not.

(2) "It is based in part on ancient and scanty inscriptions, of which there could again be multiple interpretations." Bengtson must be the first historical linguist ever to object to the use of evidence merely because it is ancient: most historical linguists, I think, would take the view that ancient evidence is a priceless commodity. In any case, Bengtson is wrong when he asserts that Michelena's reconstruction is partly based on the Aquitanian evidence. It is not: it is based upon the modern forms, including all the dialect variants, upon the medieval materials, and upon the treatment of loan words, and the Aquitanian materials are invoked only as confirmation of conclusions reached on *independent* grounds. This should be obvious to anyone who has read Michelena's book. For example, in his reconstruction of the word-initial plosives of pre-Basque, on pp. 238-244, an analysis to which Bengtson takes particular exception, there is not a single mention of any Aquitanian word.

(3) "It ignores the testimony of modern dialects." I have never seen such an outrageous misstatement in all my life. The testimony of the modern dialects, being by far the most abundant source of information and the only material to which the comparative method can be applied, is invoked *ceaselessly* by Michelena: just open the book at random and read any page. Where does Bengtson think those 4500 Basque words in the index come from? What Bengtson means by his bizarre statement, it quickly becomes apparent, is merely that Michelena's reconstructed pre-Basque phonology is *different* from the phonology of the modern language — and that's deeply inconvenient for Bengtson, because he is working exclusively with modern Basque forms. All of his comparisons involve strictly the modern forms, in defiance of the well-understood phonological prehistory of the language, and he therefore wants us to change our reconstruction merely so that his comparisons will look better. But this is absurd: the whole point of Michelena's reconstruction, and the chief reason that it is universally accepted, is that it *successfully accounts* for the modern dialect forms. What more could anyone reasonably require of a reconstruction?

Bengtson goes on to complain that Michelena's reconstruction of just 16 consonants (with the whole set appearing only intervocalically) must be wrong because modern varieties have 17-23 consonants. Why on earth is this supposed to be an argument? Early Old English had only 16 consonants; modern English has 24. Should we therefore reject the standard view of the phonology of Old English? Anyway, Bengtson is counting wrong. Michelena's inventory of 16 consonants excludes the five or six palatal consonants, even though these were probably already present in pre-Basque, because these segments had a special status and never appeared in lexical items, but only in "expressive" variants of these. But Bengtson's figure of 17-23 modern consonants *includes* the palatals. If the palatals are excluded, the modern central dialects have just 17 consonants: of the original 16, the contrasts between */L/ and */l/ and between */N/ and */n/ were lost, reducing the system by two segments, while the new segments /m/, /f/, and /j/ were added, producing the modern 17. Not much of a change for 2000 years.

(4) "It is not supported by any external comparison, and in fact external comparison refutes most of it." True, our reconstruction is not supported by external comparison (though it *is* supported by the Aquitanian data, which I consider far more important!), but so what? What principle of reconstruction obliges us to twist our reconstruction of one language to make it look like some other language not known to be related to it, just so that we can then claim that we have evidence for a genetic link? Bengtson is here openly admitting, in his own way, the truth of the central point which my entire paper was written to demonstrate: **the Dene-Caucasian comparisons are irretrievably in conflict with the established facts of Basque.** My response is the only one I consider possible: the comparisons are wrong and must be rejected. Bengtson, amazingly, wants to deny the facts in order to cling to his hypothesis. Readers can decide for themselves which of us is behaving rationally.

One aspect of Michelena's reconstruction which particularly troubles some of the readers is the absence of an /m/ in pre-Basque, which they regard as typologically unusual and therefore suspect. Unusual it certainly is, but not *that* unusual: in the famous UPSID sample of languages described in Maddieson (1984), /m/ is either totally absent or only very marginally present in 5% of the languages in the sample. In any case, typological arguments must yield pride of place to hard evidence, and Michelena could find no evidence at all for the existence of an /m/ in pre-Basque. As asserted in my paper, Basque /m/ in native words derives either from */b/ or from the cluster */nb/.

Observe that virtually every single ancient word beginning with /m/ today contains, or formerly contained, an /n/, whose presence conditioned the development of initial /b/ to /m/: *mendi* 'mountain' (which may be a loan word), *mihi ~ min* 'tongue' (< **bini*), *mahats* 'grapes' (< **banats*), *min* 'pain', *mahai(n)* 'table' (< **banane*), *mehe* 'slender' (< **bene*), *men ~ mende* 'authority, dominion', *mende* 'century', and so on. The exceptional *maite* 'beloved' is generally thought to be a loan from Celtic (Gorrochategui 1987). A few words in *ma-* are clearly "expressive" variants: *madari* 'pear' (*udare ~ udari* 'pear'), *magal* 'wing' (*hegal* 'wing'), *mako* 'hook' (*gako* 'hook', almost certainly a loan word), and a few others. Other words in *m-* are either loan words or expressive formations of no great antiquity, such as the large group of (mostly localized) adjectives denoting defects: *makal ~ mazkal* 'weak, feeble, sick', *makur* 'crooked, twisted', *murritz* 'stunted, short', *muker* 'hard; unfriendly', *motel* 'tasteless, insipid', *matzer* 'deformed, defective', *motzor* 'crude', *mokor* 'shy, unsociable; perverse', *mazkor* 'hollow, lacking the normal contents' (*e.g.*, of nuts), *mozkor* 'drunk', and a number of others.

Nor are things any different with medial /m/. There is abundant evidence that this derives variously from the cluster /nb/ (*seme* 'son' < **senbe*, attested in Aquitanian), from /b/ when an /n/ is nearby (*zamau* 'tablecloth' < Lat SABANU), or rarely from /n/ next to /u/ (*kuma* 'cradle' < Lat CUNA). Given such evidence, there is no reason to doubt that the few remaining words like *amets* 'dream' and *hamar* 'ten' have the same sort of origin. The only word for which I think an exception might be made is *ama* 'mother', but the special status of this nursery word hardly needs to be pointed out. On the other hand, we cannot dismiss the possibility that *ama* derives from earlier **anba*. Not only would this be phonologically perfect, but it would suggest a combination of the kinship element **an-* posited by Michelena as present in *anaia* 'brother (of a man)' (< *? **an-na-*), *ahizpa* 'sister' (< **anizpa*), and *ahaide* 'relative' (< **anaide*) with the morph *-ba* which is so frequent in kinship terms.

Note further that /m/ is absolutely lacking from the inflectional morphology of both verbs and noun phrases: most of the other Basque consonants make an appearance somewhere in the rich inflectional morphology of the language, and /n/ in particular is very frequent indeed, but there is no trace of /m/ — because it wasn't in the language when most of the inflectional morphology was being constructed. Nor do we find a single instance of /m/ in grammatical words like conjunctions, postpositions, and pronouns. Nor do we find /m/ in the very numerous word-forming suffixes of Basque, save only in the noun-forming suffix *-men ~ -mendu*, a transparent loan from Latin -MENTU. All this is highly embarrassing for those who would posit an /m/ for pre-Basque merely in order to make their comparisons go through.

In short, there is no evidence for an /m/ in pre-Basque, and assertions to the contrary constitute nothing but wishful thinking. Comparisons which require a pre-Basque /m/ may be dismissed out of hand.

On the other hand, I have been persuaded by Hualde that my assertion that no native Basque word can begin with a voiceless plosive must be modified. For /p/ and /t/, my statement remains correct: no ancient Basque word begins with either of these, and it is not even certain that the phoneme /p/ was present at all in pre-Basque — if it was, it was certainly very rare. But there do appear to be a very few words with initial /k/ whose antiquity seems assured, or at least plausible: *ke* 'smoke', *koipe* 'grease, fat', *kirats ~ kirets ~ karats* 'stench, foul-smelling; bitter' (cited by Bengtson in item [197]), and the doubly anomalous *kalte* 'harm, injury', with its illicit *-lt-* cluster, and possibly one or two others. In addition a few other words which normally have voiced initials exhibit voiceless ones in severely localized areas: *gar* 'flame' is *kar* in places, and the compounds *gurpil* 'cartwheel' and *bikain* 'excellent' appear as *kurpil* and *pikain* in a few localities. Michelena (1977:242) attributes this phenomenon to voicing assimilation induced by the voiceless second plosive, and he includes in this analysis the case of *kalte* (also attested as *galte*, and apparently derived from *galdu* 'lose'; Michelena proposes **galdu-te*, which would at once yield the attested *galte*). The same might be true of *koipe*. For *kirats,* this explanation is not available, but the variation in form, plus the existence of some odd additional meanings listed in Azkue's dictionary, leads me to conclude that this is merely another "expressive" formation of no great antiquity and hence that it can be safely discarded (see section 5 below). (In fact, there is another explanation: G has a word *keats* 'stench', a compound of *ke* 'smoke' and *ats* 'stench', and the other words might in principle be developments of this with an *r* inserted to break up the hiatus; this is not obviously right, however, since such use of *r* is not common in Basque, though it is certainly attested.) The monosyllabic nature of *ke* and *kar* is probably significant: in Michelena's reconstruction, monosyllables alone would have had initial stress, possibly favoring a voiceless realization of the initial plosive. (Recall that a crucial feature of pre-Basque phonology was the absence of contrasts like *k/g* in initial position.)

But that's it. The existence of a handful of irregular forms, at least some of which have plausible explanations, does not in the least justify the long-rangers in their practice of assuming that any convenient Basque word with initial /k/, or still less /p/ or /t/, can be safely projected back to the remotest period of the language. It is by no means rare for a language to exhibit two or three exceptions to an otherwise regular phonological development. Consider the treatment of the Latin initial clusters PL-, CL- and FL- in Castilian Spanish. These regularly developed into a palatal lateral; hence, for example, PLANU > *llano* 'flat'; CLAMARE > *llamar* 'call'; FLAMMA > *llama* 'flame'. But three or four words are exceptional, such as PLATEA > *plaza* 'town square' and CLAVICULA > *clavija* 'peg' (Penny 1991:62-63). These are not learnèd forms; they are just exceptions. No one is proposing that their existence requires a major rethinking of the phonological history of Spanish.

Bengtson queries Michelena's (and my) view of the pre-Basque plosive system, proposing instead a pre-Basque contrast between unaspirated /b d g/ and aspirated /ph th kh/ in word-initial position. Well, this suits his purposes, but of course it runs into difficulty with the Basque data. In particular, why is it that, in modern Basque, initial /b g/ are so exceedingly common in all parts of the vocabulary, while initial /p t k/ and initial /d/ are almost entirely confined to words which everyone (even Bengtson) must admit are loan words? Where are the numerals, the pronouns, the grammatical words, the body-part names, the names of indigenous plants, the simple adjectives, with initial /p t k d/? Why is it that words with these four initials are almost always shared with Romance, whereas words with initial /b g/ usually are not? How does Bengtson explain this?

Bengtson also takes exception to my remarks about the Basque aspiration, which, following Michelena, I see as normally suprasegmental in origin, and which I see as having been extended in Z to all monosyllables which can bear it. He makes three points: (1) Basque /h/ must be ancient, because it is attested in Aquitanian; (2) Z has a few monosyllables lacking the aspiration; (3) the aspiration must be segmental in origin to make his comparisons work. First, nobody has ever denied that Basque /h/ is *ancient*, and the Aquitanian /h/'s in fact *confirm* Michelena's reconstruction. Second, Bengtson's three examples of Z monosyllables without aspiration can all be disposed of. Z *ar* 'male' represents an item which most commonly occurs as the second element in a compound, where it would usually not be aspirated, and it is very common in Basque for the combining form of such elements to be generalized as independent words; see the discussion in Michelena (1957-58) or in Trask (forthcoming c). Z *ots* 'male animal' represents a word which is otherwise *orots* (or even *ordots*), and *ots* is nothing but a localized contraction of *orots*, showing the usual Z loss of intervocalic /r/: the word is not a monosyllable. And Z *ütz* 'leave' is merely the stem of the verb *ützi* (common *utzi),* and verb stems are not subject to the extension of the aspiration. Moreover, LN Sal old G have *eutzi* for this verb, B has *etxi*, and even old Z has *eitzi* (< **eutzi*), demonstrating that **eutzi* is the original form of the verb, and that even the stem was not originally a monosyllable (Michelena 1958, 1977:493). Third, Bengtson cannot argue against the received view of Basque */h/* on the ground that it's inconsistent with his comparisons when it is the validity of those comparisons which is at issue in the first place: this is the man who accuses me of "illogical circularity".

Just to complicate things, Hualde reminds me that the demonstratives *hau*, *hori*, *hura* appear in R and Sal as *kau*, *kori*, *kura* and in Aezk as *gau*, *gori*, *gura*, a fact which leads Michelena to conclude that these items must once have had initial plosives in all varieties. If he is right, then at least these three items contain an etymological /h/, but these items are isolated and enigmatic: nothing else in the language shows this kind of variation.

There is a further point about the phonology of pre-Basque, one which I did not pursue in my paper, but which presents yet further difficulties for many of the proposed Basque-Dene-Caucasian comparisons. This is the phonotactics, or, if you prefer, the morpheme-structure rules. Native Basque words overwhelmingly conform to certain very restrictive patterns of segments. This is a topic which has so far been little investigated (though see Trask forthcoming *c*, *d* for some discussion). But I want to draw attention to one particular fact here: in a Basque word of the form CVCV-, it is very rare for both consonants to be plosives (Lakarra 1995). When they are, the first plosive is, of course, always voiced, and the second one is almost always voiced. We have the following words with two voiced plosives: *begi* 'eye', *bide* 'road', *bat* 'one' < **bada* or **bade*, *biga* 'two', *bider* 'time, occasion' (Fr *fois*), *gabe* ~ *bage* 'without', *gogo* 'mind, soul' (possibly a reduplication), *gudu* 'combat', and *gogor* 'hard' (almost certainly a reduplication of *gor* 'deaf', but formerly 'unyielding'). With a voiceless plosive, we find only *beti* 'always' and *guti* 'not much', both of which probably involve the ancient suffix *-ti*, *bete* 'full', which as Zabaltza remarks is often thought to be related to *bat* 'one', and the exceptional *tipi* ~ *tiki* 'small', whose initial /t/ marks it indisputably as an expressive formation. There are no others.

Consequently, even allowing for the possibility of */b/ >/m/, the presence in remote comparisons of such Basque words as *beko ~ moko* 'forehead; beak', *bekoki* 'forehead', *kaiku* 'wooden bowl', *kako ~ gako* 'hook', *kokot* 'nape', *matel* 'cheek', *moto* 'headscarf', *mutur* 'snout, muzzle', *pataxa* 'bottle', *pipil* 'bud', and *potorro* 'vulva' must, quite apart from their other very severe problems, be regarded with the deepest suspicion. Such words just do not look like native Basque words, and any comparison which depends crucially upon the use of such words is an edifice built upon sand.

Finally, Bengtson offers some provisional attempts at a reconstructed "Macro-Caucasian" phonology. Naturally, I applaud this, since only such work, successfully performed, can demonstrate that we are looking at something more than chance resemblances. On offer here is a set of "correspondences" involving the Caucasic lateral affricates. His nine examples involve the following matches; Basque reconstructions are his, and I can't tell if Caucasic /w/ is supposed to be significant or not:

	Basque	Caucasic
1.	-rd-	dl-
2.	-d- < *-rd-	tl(w)-
3.	-d- < *-rd-	-tl'-
4.	-rd-	-tl'-
5.	-rd-	-rtl'-
6.	l-	tl'-
7.	l-	-ldl(w)-
8.	-r	-rtl'-
9.	-l	-(w)tl -

I note with interest that no two of the nine comparisons involve the same pair of segments, which is disappointing, the more so since the remaining segments in the nine matches show no discernible pattern. Anyway, Basque item 1 involves a (misglossed) bimorphemic word; 2 involves an *ad hoc* reconstruction; 3 involves an obviously secondary dialect variant (and a reconstruction which is clearly wrong); and 6 involves a loan from Latin.

4. The segmentations. One aspect of the comparisons that I have objected to particularly often and particularly strenuously is the constant practice of all the authors of introducing arbitrary segmentations into the words they compare, in order to remove from consideration any segments or syllables which they cannot match. With just one significant exception, discussed below, this is done in a totally unprincipled manner: the portions that are segmented into oblivion are neither identifiable nor recurrent, and they are removed for no better reason than that they are not convenient. I find this practice reprehensible, and I am not alone in my view: in all his various commentaries upon the efforts of earlier comparativists, Michelena is equally scathing about the same practice. No doubt my critics will see my position as yet one more piece of bloody-mindedness on the part of a blinkered and reactionary traditionalist who cannot open his eyes to modern methods. So let me ask this question: if the comparativists I am criticizing really were just arbitrarily chopping off inconvenient bits of words to make invalid comparisons look better, how would the result look different from what we see here?

Let's consider first the apparent "suffixes" which our authors remove. Excluding the cases in which the morph removed genuinely is a separate morpheme, I find that the authors, explicitly or implicitly, variously remove from their Basque words final *-ar* (items [166], [294], [297], [304], [308]), *-atx* ([197]), *-bin* ([264]), *-bor* ([307]), *-da* ([264]), *-dur* ([264]), *-el* ([217], [251]), *-en* ([310]), *-ga* ([107]), *-gar* ([107]), *-hezi* ([60]), *-i* ([131]), *-il* ([281]), *-in* ([222], [244]), *-io* ([146]), *-kho ~ -go* ([298], [307]), *-l* ([209]), *-leta* ([273]), *-lo* ([201]), *-n* ([82], [91], [124], [175], [185], [186]), *-ne* [97]), *-no* ([216]), *-o* ([184], [217]), *-on* ([116]), *-pil* ([125]), *-r* ([69], [75], [77], [87], [128], [180], [206], [228], [229], [302]), *-ri* ([210]), *-rro* ([247]), *-sa* ([243]), *-ter* ([205]), *-ton*, *-tan(a)* ([259]), *-tz* ([159]), *-tzar* ([106]), *-tzu* ([311]), *-tzurrin* ([114]), *-u* ([125]), *-un* ([58], [203]). In no single case do they make the slightest attempt to identify the obliterated morph, or to justify its removal in any way beyond the obvious fact that it doesn't match. In other cases they appear to remove a valid final morpheme purely by accident, as in the infamous item [211]. In still other cases, the proposed matchup is so faint that I can't tell just what is being removed, as in

item [289]. And I haven't even looked at the cases in which final morphs are chopped off words in the other languages.

With the alleged "prefixes", things are no better. We find the following initial morphs removed from Basque words: *a-* ([2], [4], [5], [7], [8], [10], [11], [12], [13], [14], [19], [20], [28], [29], [30], [32], [34], [35], [37], [130]), *az-* ([34]), *be-* ([51], [54], [56], [57], [58], [134]), *beha-* ([54]), *bi-* ([70], [72], [74], [75], [78]), *e-* ([83], [84], [86], [87], [98], [100], [101]), *ema-* ([88], [89]), *ha-* ([127]), *han-* ([129]), *he-* ([140]), *i-* ([148], [156], [158], [159], [170], [172], [174], [176], [178]), *in-* ([166]), *ma-* ([211]), *o-* ([230], [232]), *oi-* ([233]), *u-* ([159], [280], [281], [282], [283], [288]), *zu-* ([313]). This time, of course, there are a few efforts at identifying *some* of these "prefixes". Two of the authors try to justify *be-* and *bi-* as "body-part prefixes", but of course, as my paper points out, they only remove these "prefixes" when they find it convenient to do so, retaining them instead when *that* is more convenient, and Bengtson even removes this "prefix" from *bizi* 'life', which is not a body-part name nor even a concrete noun. Bengtson's further interesting proposal of a set of fossilized "noun-class prefixes" I have already dealt with, and his reply here adds little of substance.

De Grolier takes exception to my dismissal of these ghostly "prefixes", on the decidedly peculiar ground that prefixes are well attested in Indo-European, Sino-Tibetan, and Khoisan. No doubt this is true, but it is hardly relevant. The point is that, except in verbal morphology, ancient prefixes are *not* attested in Basque. Anyone wishing to make a case for fossilized prefixes in the language is therefore obliged to offer something in the way of hard evidence for such a claim. So far we have seen no such evidence, but only vague suggestions, unsubstantiated assertions, and circular reasoning: any arbitrary sequence that doesn't match is declared to be a "prefix" in order to force the remaining material to match something, and the fact that some kind of match is finally obtained is invoked as a demonstration that the morph removed must indeed be a prefix.

The same reader hypothesizes that, since I object to such unsubstantiated assertions in the case of Basque, I would presumably also automatically deny Meillet's recognition of certain prefixes in Indo-European, a family in which prefixes are abundant. I confess I am unable to follow his reasoning here: why does he think that such an absurd suggestion is a useful contribution to the discussion? He goes on to point out that Schuchardt and Berger had previously suggested the presence of certain prefixes in Basque. But he himself adds at once that Berger has withdrawn his suggestion, and, as for Schuchardt, his suggested prefixes have proved to be without foundation and are recognized by no one today. These precedents do not appear to be helpful to the current proponents of yet further prefixes.

De Grolier further declares that I apparently require a candidate prefix to appear in every available word before it can be accepted as a true prefix. I see what he means, and I guess I should have made myself clearer. The reference is to one of my several reasons for rejecting the "body-part prefix" *b(e)-* as unsubstantiated. The point I was trying to make is that this alleged prefix shows no discernible distribution. All we have is the banal observation that a minority of body-part names begin with *be-* or *bi-* or at least *b-*. Unlike some of their predecessors, the present authors do not even attempt to confine their "prefix" to names of paired body-parts: they happily remove it from the words for 'forehead', 'back', 'heart' and 'bile'. Since, as my paper makes clear, initial *b-* is so exceedingly common in Basque anyway, I cannot see that anyone has made even the beginnings of a case for such a prefix: our authors declare this morph a prefix merely because they find it convenient to remove it from consideration, except of course in those words from which its removal is not convenient, and in which it is therefore declared not to be a prefix this time.

Finally, I remind the reader that, in almost every single case of a Basque word whose morphological structure is clear, such as the obvious compounds and derivatives, the authors' segmentations are plainly wrong. If they get the clear cases wrong, is it likely that they've got the opaque cases right?

5. The phonesthetic items. Another practice I have repeatedly objected to is the repeated use of phonesthetic, or "expressive", formations in drawing comparisons. In Basque, these are of two types: (1) expressive variants of ordinary lexical items, and (2) phonesthetic items coined from scratch. As for the first type, I have nothing to add beyond what I say in my paper: anyone who does not believe me when I assert, for example, that native lexical items do not contain the segment *tx-* is invited to read the evidence presented in vast detail in Michelena (1977) and in several others of Michelena's works.

On the second type, de Grolier again objects to my dismissal of these items as comparanda, on the ground that many eminent comparativists have willingly made use of similar forms in connection with other languages (Indo-European, Dravidian, Uralic). True, but there are two vital differences. First, these other scholars were working on languages whose genetic affiliation *had already been established.* They did not appeal crucially to such items to establish the genetic link in the first place. But the present case is utterly different: no genetic link has been established between Basque and any of the other languages under discussion, and making crucial appeal to the notoriously unreliable phonesthetic items in order to try to establish the required links is simply unacceptable. *Once* somebody has established beyond reasonable doubt that Basque is related to something else, we can then look to see if any phonesthetic items have survived from the ancestral language.

Second, there is overwhelming evidence that the Basque items in the second class do not continue ancient formations but are instead of very recent origin. This evidence is as follows: (a) these items regularly have phonological characteristics which set them apart from indisputably ancient words, such as initial voiceless plosives, the frequent presence of /m/ and of the palatal segments, and extensive reduplication; (b) a number of these items exhibit peculiar local variations in form, variations which in no way parallel the ordinary phonological differences among regional varieties; (c) very often each such word displays a startling range of unrelated senses; (d) scarcely a single such item is found throughout the Basque-speaking region; instead, each one is found in a more or less severely localized area, with notable regional preferences in their forms.

Consider, for example, item [244], Basque *pinpirin* 'butterfly', compared by Bengtson with various insect and spider names in other languages and cited by Bengtson and Ruhlen as continuing the alleged Proto-World root **PAR 'fly'. First, this word could hardly look less like a ancient lexical item, with its initial /p/, its reduplication, and its lack of voicing of the second plosive after /n/. Second, the word exhibits the variants *pinpirin*, *pinpirina*, *pinpin*, *pinpina*, *pinpiro*, *pinpirineta*, and *pinpilinpauxa* (at least), to which Bengtson adds *pinpilin,* not listed in any of my dictionaries. Third, the word means not only 'butterfly' but also 'bud' (of a flower), 'garfish', 'undeveloped fruit', 'elegant', and 'pretentious, putting on airs'; its reduplication *pinpili-panpala* means 'favorite'; its diminutive *pinpirinatxa* means 'little girl decked out in her finery'. Fourth, the word is confined to a small area, essentially to the province of Lapurdi and a few towns in the neighboring provinces; other parts of the country have quite different words for 'butterfly', such as *mitxeleta*, *tximeleta*, or *inguma.* And, as pointed out in my paper, this region strongly favors expressive formations in *pin-* or *pan-*.

So: which is more likely? That *pinpirin*, which looks just like one more of the many hundreds of recent and severely localized expressive formations in Basque, is exactly that? Or that it amazingly continues a 7000-year-old root with an utterly un-Basque form?

In connection with several items, such as [62], involving Basque *bero* 'hot', and [212-213], involving the nursery word *mama,* de Grolier objects to my rejection of such forms as too universal to be of any use in comparisons. He seems to believe that this position is a weasly cop-out of my own invention. Not so: such words are universally excluded from comparisons by historical linguists because they are so treacherous as to be devoid of value. See, for example, Anthony Fox's recent book on the comparative method (Fox 1995), in which he declares that items like *mama* "should...be excluded from the process of comparison" (p. 63), or see H. H. Hock's admirable textbook (Hock 1986), in which, on page 559, he makes exactly the same point about imitative and nursery words. "Mainstream" historical linguists like Fox and Hock adopt this cautious policy, not because they are eager to plant their establishment jackboots firmly on the necks of pioneering long-ranger upstarts, but merely because 200 years of painful experience has demonstrated beyond any debate that such a policy is indispensable. Those who reject such sound advice are doomed to find spurious connections wherever they look.

As for item [269], involving Basque *tu* 'spit' and similar words in other languages, de Grolier protests at my dismissal of this as being obviously imitative in origin. So, let me ask this question: if you were going to invent an imitative word for 'spit', what would you come up with? I think it would be very difficult to improve on the form *tu*. The speakers of Basque, Burushaski, and the other languages cited apparently share my opinion of the excellence of such a choice.

6. The data from the other languages. In my paper, I stress that I am competent only to judge the use of Basque materials, but I ask whether specialists in the other languages involved would find fewer errors in the data adduced from those languages than I have found in the Basque materials. I now have my answer. For Sumerian, de

Grolier points out that the long-rangers have used forms which are only attested very late in that language and which cannot safely be attributed to the earlier language, and he also complains that they have failed to recognize the frequent alternations in the language, which "in rather numerous cases" would require different reconstructions from those assumed by our authors. It would appear, then, that, with Sumerian as with Basque, the comparativists have arbitrarily selected any forms which they find convenient, without regard to what is known of the history of the language.

For Burushaski, things are even worse. De Grolier complains that our authors have used only obsolete and unreliable sources of information, as a result of which they have frequently included loan words from several neighboring languages. For Burushaski (and also for Sumerian), de Grolier informs us that he could provide "a long list of errors or misinterpretations" on the part of Bengtson and his colleagues; he considers, unfortunately, that he lacks the space to do this.

De Grolier's misgivings are confirmed by Tiffou, who also expresses surprise at the dubious quality of the Burushaski data cited. He finds space to examine and criticize only a small selection of the comparisons, but his conclusions are unambiguous: many of the comparisons are spurious or erroneous for various reasons (loan words, wrong transcriptions, wrong segmentation, wrong analyses, non-existent forms cited), and "most of the analyses can be criticized" from the Burushaski side just as I have done from the Basque side. Even though Tiffou concludes on a positive note, accepting a few of the proposed comparisons as plausible, this is damning stuff, and a number of the comparisons which I could not fault on the Basque side are now seen to be unsustainable on the Burushaski side. The case for "Macro-Caucasian" has fallen apart, and with it a large chunk of the case for "Dane-Caucasian".

In my paper, I decline to repeat the data from Sino-Tibetan and Na-Dene languages, a policy which is openly declared. Ruhlen takes me to task for this policy, asserting that it has disastrous consequences for my case.

I omitted the data from these other languages for two reasons. First, including it would take up a great deal of space, and hence it would add perhaps another twenty pages or so to what is already a distressingly long paper, and there must be a limit even to the good will of the editors of *Mother Tongue.* Second, and far more importantly, Ruhlen is in error when he asserts that the additional comparisons are of crucial importance. He seems to believe that I am trying to criticize the entire Dene-Caucasian edifice, but this is not so. As my paper makes perfectly clear, I am criticizing *only* the use of Basque materials and the consequent attempt to include Basque in this hypothetical macro-family. In most cases, my criticisms take one of the following forms: (1) the Basque word was not in the language as recently as 2000 years ago; (2) the Basque word can be shown to have formerly had a very different form, one rendering the comparison untenable; (3) the Basque word has been wrongly analyzed, or at least arbitrarily segmented to an unacceptable degree; (4) the Basque word is a phonesthetic word or a nursery word and should not, as a matter of principle, be included in a comparison; (5) the Basque word does not have the meaning imputed to it; (6) the alleged Basque word does not exist. If I can remove the *Basque* word from the comparison in one of these ways, therefore, the existence of any number of Sino-Tibetan or Na-Dene words is utterly irrelevant.

A fine example of this is Basque *gorotz* 'dung', which is explicitly singled out by Ruhlen as a case in which the data from Sino-Tibetan and Na-Dene would prop up the comparison. But, as I point out in my original paper and in section 2 of this response, it is hardly likely that *gorotz* is a native Basque word. And, if the Basque word is borrowed from Romance, who cares what the Tibetan or Apache words for 'dung' might look like?

Curiously, Ruhlen also chides me for ignoring certain work by Starostin and by Nikolaev on Dene-Caucasian. But these papers pay no attention to Basque, and hence they are quite irrelevant to my purpose.

There is a further point here, and a very important one. Like many long-rangers, Ruhlen appears to believe that increasing the number of languages in a comparison somehow reduces the frequency of chance resemblances and correspondingly increases the degree of confidence we can place in apparently positive results. But this is a shrieking fallacy, and one which has been thoroughly demolished by Don Ringe (1992). Ringe is able to prove (not argue, *prove*) that this belief is the very opposite of the truth. Adding more languages to a comparison does *not* reduce the incidence of chance similarities; instead, by giving chance a much larger number of cases to play with, it greatly *increases* the likelihood of chance resemblances. Since Ringe's work is now very well known, I am astounded to find anyone apparently still maintaining the opposite, and totally fallacious, view.

Ruhlen himself offers a magnificent example of Ringe's demonstration. John Bengtson, in the work under review here, compares Basque *odol* 'blood' with Burushaski *del* 'oil; contents of an egg' and Sumerian *dal* 'breath,

life, soul', a comparison which I think few people will find compelling. Apparently there is nothing even remotely suitable in Caucasian or Yeniseian. Ruhlen now points to Eyak *dehl* 'blood' and to Proto-Athabaskan **dehl* 'blood', from which I assume I am meant to conclude that the Proto-Na-Dene for 'blood' was something like **dehl*. OK; I'm willing to accept that. But Ruhlen then goes on to point to the resemblance between Basque *odol* and Na-Dene **dehl*, and to conclude that an objective linguist would very likely "conclude that the entire etymology is almost certainly valid".

I find this argument simply beyond belief. Consider what is going on. Bengtson can find no match for Basque *odol* in Caucasian. In Burushaski, he can find nothing better than a phonologically similar word which means nothing like 'blood' — and, phonotactics permitting, he was highly likely to find something like that. For Sumerian, he can find only the same disappointing result. For Hattic, Hurrian, Urartean, apparently no luck. Yeniseian? Another strikeout. Sino-Tibetan? Nothing. Na-Dene? Aha! *This* time we strike it lucky, and find a word for 'blood' which looks quite a bit like the Basque word.

But look: if we keep adding more and more languages to the comparison, then sooner or later, by the ordinary laws of probability, we are *bound* to find some kind of match for any given item. This is precisely what Ringe demonstrates with mathematical rigor, but the lesson is lost on Ruhlen. For him, only confirming instances get counted, and no amount of negative evidence can persuade him that his methodology, which is perfectly designed to sweep up any number of chance resemblances, is doing precisely that — and no more.

Consider a parallel. Turkish has a word *tepe* 'hill'. Suppose that I, an enthusiastic long-ranger, note the existence of ancient Egyptian *tp* 'head' and of English *top*, and propose that these three form part of a single etymology. Not convincing so far. But then, adding a few more languages to the comparison, I stumble across Nahuatl *tepe-tl* 'hill', in which *-tl* is an absolutive suffix, not part of the stem. By Ruhlen's criteria, it seems this new evidence is so powerful that our hypothetical objective linguist will "conclude that the entire etymology is almost certainly valid". Is this any way to go about things?

In Ruhlen's case, we have a whole bunch of languages chosen in advance for comparison. Two of them happen to turn up with somewhat similar words for 'blood', and two more have somewhat similar-looking words which don't mean anything like 'blood'. Why is this interesting? What was he expecting to find?

7. My choice of words. De Grolier takes me to task for using such charged words as "preposterous", "absurd" and "outrageous" in evaluating some of the proposed comparisons, and Wescott suggests that I am in danger of "exceed[ing] the bounds of collegial civility". No doubt they are right; my language could be more tactful. But I have used such language only in a handful of places in my paper, and those are places in which I consider these adjectives to be entirely called-for. Let me consider one of these, number [211].

Here Bengtson has misanalyzed Basque *makutsik* 'in one's shirtsleeves' as containing an alleged ancient root **-kuts-* 'sleeve'. Suppose a long-ranger who knew no English were to stumble across the English word *shirtsleeves*, cheerfully remove the first and last elements, extract an alleged root **-irtsl-*, assign to this the meaning 'sleeve', and then compare this with something in Caucasian. What adjective would de Grolier consider appropriate here? "Bold"? "Unexpected"? "Questionable"? Surely any linguist would at once dismiss such an analysis as preposterous or worse. Why should the rules be different for Basque? (I note that Bengtson has now, quite properly, withdrawn this comparison.)

Incidentally, de Grolier is somewhat in error in commenting, in connection with item [83], that the "monumental confusion" I refer to derives in part from my combining different works by different authors. I realize now that I could have been clearer, but the three comparisons I single out here for criticism are consecutive entries — numbers 67, 68 and 69 — in a single work, Bengtson (1991a).

Further, de Grolier and Bengtson protest about my repeated use of the word "destroyed" in dismissing some of the comparisons on offer. Well, if demonstrating that the Basque word involved did not enter the language until post-Roman times does not constitute destruction of the comparison, what on earth *would* destroy it?

Wescott suggests that my comment on Michelena's dismissal of earlier comparative work implies that I am maintaining that long-rangers are necessarily "frivolous scholars". That was never my intention, and I apologize if I have inadvertently given such an impression. To begin with, I am commenting in my paper only on specific named publications, and not on long-range work generally. Furthermore, apart from my single suggestion (now vindicated) that the evidence adduced from other languages might very well prove to be as flawed as the Basque data, I have

commented solely and exclusively on the use of Basque data in the cited works, and on nothing else whatever. It is true that I consider the methodology adopted in the cited works to be fatally flawed. Deciding in advance what you want to prove and then proceeding merely to collect possible confirming instances, with no attention at all to any amount of counter-evidence, with no consideration of alternative explanations, and with no scrutiny of the outcome, can never, in my view, lead to any useful results at all. It is also true that I believe it virtually impossible to do useful work on languages you are not personally acquainted with, and absolutely impossible to do useful work in ignorance of, or in defiance of, firmly established findings on the history and prehistory of the languages involved. The huge number of errors I have identified surely makes that point more tellingly than any amount of argument. I am not maintaining that the long-rangers whose work I discuss are frivolous; I am only maintaining that they are wrong to draw the conclusions they draw from such inadequate methodology and such blatantly faulty data.

Finally, de Grolier chides me for "immaturity" in employing such vocabulary, asserting that such emotive language has long since been abandoned in the apparently more mature natural sciences. I can only assume he is not very well-read in the natural sciences; I would find it easy to provide a list of books detailing recent stormy professional disputes among even the most distinguished specialists in such varied fields as cosmology, paleoanthropology, and evolutionary theory, disputes which have often involved abusive personal attacks far, far beyond my use of "preposterous" or "outrageous", and which have at times led the warring scientists to stop speaking to one another and even, it would appear, to try to block the publication of one another's work. Historical linguistics, even as practiced by me, is a sedate and gentlemanly pastime by comparison.

8. The Dene-Caucasian hypothesis. While I have not concerned myself with the Dene-Caucasian hypothesis as a whole, Ruhlen has chosen to make an issue of it in his response, and he raises several points to which I feel obliged to reply.

First, Ruhlen appears to believe that the Dene-Caucasian hypothesis must be taken seriously because it has been asserted and because its proponents have adduced hundreds of comparisons among the languages they assign to the grouping. He objects to my characterization of the collection of languages compared with Basque as an essentially *a priori* one, and suggests that, if I am right, I should be able to produce a comparable body of "evidence" for any arbitrary collection of six language families involving Basque.

Yes, I agree: if I am right, then I should indeed be able to do this. Moreover, I am very confident that I *could* do it, if I were so inclined. After all, no one who has ever tried to compare Basque with anything at all has failed to find "evidence" in the form of chance resemblances, and the length of each such list appears to reflect nothing more than the amount of time and effort devoted to compiling it. But, as it happens, I am quite unwilling to devote five or ten years of my life to pursuing an empty exercise in order to make a point. Nor am I being in any way unreasonable in this.

The onus of proof is not on me to demonstrate that the Dene-Caucasian hypothesis is wrong; rather, it is on the proponents of Dene-Caucasian, or of any other ambitious proposal, to demonstrate that their accumulation of comparisons constitutes something more than butterfly-collecting, something more than the accumulation of the abundant chance resemblances which can always be found in such enterprises. At the very least, they need to attempt some scrutiny of their results — for example, by applying the statistical tests described in my next section. But they do nothing.

Indeed, I have the clear impression that the proponents of Dene-Caucasian believe that their hypothesis is greater than the sum of its parts — that is, that the hypothesis is so compelling that no amount of criticism of the details can be allowed to call the hypothesis into question. For example, I note with interest that, in spite of my telling criticisms, not one of the proponents of Dene-Caucasian has even briefly entertained the idea that the inclusion of Basque in Dene-Caucasian might have been a mistake, even if the hypothesis is otherwise tenable. Apparently Basque just has to be there, and there's no more to be said.

Amazingly, Ruhlen goes so far as to suggest that the place of Basque in Dene-Caucasian must be secure because this is the *only* "audacious" proposal linking Basque to five other families which anyone has so far advanced. In other words, we vasconists have to play the Dene-Caucasian game because it's the only game in town. I am rendered speechless by such argumentation.

A far more important point is Ruhlen's insistence that the Dene-Caucasian hypothesis must be right because "many aspects of Basque can only be understood in terms of the Dene-Caucasian hypothesis". But this is false, blatantly, howlingly false. To date, every single advance in understanding the prehistory of Basque, without exception, has come from the work of vasconists who ignore the Dene-Caucasian hypothesis. The Dene-Caucasian hypothesis has succeeded in shedding *no light whatever* on any single problem in Basque prehistory. The only attempt known to me to shed such light is Bengtson's proposal that the frequent initial vowels of Basque might be interpreted as fossilized noun-class prefixes. Had Bengtson managed to make this work, I would agree at once that Dene-Caucasian had made a contribution to my field. But he has not made it work: intriguing as the idea is, it is so far a conspicuous failure. The contribution of Dene-Caucasian to Basque historical linguistics has so far amounted to precisely zero: there is *no* aspect of Basque which can be understood in terms of the Dene-Caucasian hypothesis, nor, in spite of Ruhlen's resoundingly confident final paragraph, can I see the slightest reason to expect that this situation will change in the future. (On all this, see now Trask forthcoming *d.*)

Perhaps I should say something about Bengtson's "noun-class prefixes", since he returns to the issue in his comments. Most North Caucasian languages have noun classes marked by prefixes; Diakonoff and Starostin have reconstructed the singular class prefixes for the four classes they recognize as **u-*, **j-*, **w-*, and **r-*. Fine, but what has this to do with Basque? Well, Bengtson informs us, Basque has some words beginning with *u-* or *o-*, some words beginning with *i-* or *e-*, some words beginning with *bi-* or *be-*, and some words beginning with *ar-*. So it does. So do most of the languages on the planet, including English. Why is this an interesting observation? What has been explained?

Ruhlen argues further that Dene-Caucasian etymologies for Basque words are to be preferred to more conventional ones, and in particular to loans from Romance — indeed, he speaks of "mysterious borrowings from unidentified Romance sources" as though I were up to something rather fishy. But I'm not. Take the case of Basque *moto* 'headscarf' (and other senses), which I indeed regard as a loan from an unidentified Romance source. The point is that this word just doesn't *look* like a native Basque word: it looks for all the world like a loan word, and it must surely be one, even if we can't find an obvious source for it. But Ruhlen appears to have no interest in whether *moto* looks like a native Basque word or not: all that matters to him is whether it looks like something in Burushaski.

Does this make any sense at all? One the one hand, no single instance of a Basque cognate in Burushaski or Caucasian has been established to general satisfaction. On the other, Basque has indisputably borrowed thousands of words from Latin and Romance; for most of these we can find obvious sources, but once in a while we can't. So what? We can't find a source for English *boy*, either, even though we are certain, from its form, that it must be a loan word. Would Ruhlen therefore conclude that *boy* must be cognate with something in Burushaski?

Finally, Ruhlen reels off the names of a number of linguists who, he says, have advocated the Dene-Caucasian hypothesis "in various shapes" and declares that I would maintain that "all these scholars have merely deluded themselves into seeing historical relationships where none exist". Well, I could take issue with his list insofar as it pertains to Basque: Trombetti and Swadesh in fact compared Basque with pretty much everything on the planet; Bouda compared it with North Caucasian, Kartvelian, Uralic, Munda, Austronesian, Thai, and Chuckchi-Kamchatkan, which doesn't look to me like anybody's version of Dene-Caucasian; and Sapir never mentioned it at all. However, since Ruhlen raises the point, I might as well come clean and admit that, yes, his description is a fairly accurate one, though I would put it a little more carefully myself: so far as Basque is concerned, I believe that all these scholars have deluded themselves into seeing *evidence* of relationships where no such evidence exists. Indeed, I would maintain that the very success of Trombetti, Swadesh, Bouda, Bengtson, and the rest in finding "cognates" for Basque wherever they have looked is the strongest possible evidence that this whole methodology is untenable.

9. The remaining comparisons. I have not attempted to prove that Basque is not related to Caucasian, or to Burushaski, or to Sumerian, or to anything else. That would be an impossible undertaking. I have only attempted to demonstrate that the evidence adduced in support of such genetic links does not stand up.

Several of the readers have responded to my criticisms more or less as follows: "In pioneering work of this sort, it is inevitable that some proposed comparisons will prove to be wrong. But the identification of such errors should not be taken as invalidating an entire proposal of genetic relationship so long as other comparisons remain

secure." Of course, I agree with this, but this methodological principle is not the point here. The point is rather this: if *most* of the proposed comparisons can be shown to be wrong, is there any reason to accept the proposed genetic relationship as vindicated?

I believe I have succeeded in demolishing the majority of the "Dene-Caucasian" comparisons involving Basque. (Bengtson reckons I have disposed of only 73, but I make it well over 150.) The testimony of Tiffou and de Grolier shows that an unspecified number of further comparisons can be dismissed with a knowledge of Burushaski and Sumerian. But there remain some dozens of comparisons which I have been unable to fault from the Basque side. Or, rather, let me put that more precisely: there is no reason to doubt that the Basque words adduced were in the language at least 2000 years ago and that they had approximately the forms and meanings imputed to them. But does this mean anything?

Certainly not. All that we have left, after the removal of a huge number of unsustainable comparisons involving Basque, is some dozens of vague chance resemblances between some Basque words and some words in the other languages. But such chance resemblances can always be found, and they count for nothing. Don Ringe (1992) has recently demonstrated, with merciless mathematics, that chance resemblances between arbitrary languages are far commoner than our naïve expectations would lead us to suspect, but the linguists whose work I am criticizing pay no attention. We now have simple and reliable statistical methods for judging whether some collection of resemblances is more substantial than we would expect from chance alone (Oswalt 1991, Ringe 1992), but the proponents of Dene-Caucasian have shown no interest in such rigorous methodology. (If someone can provide me with a Swadesh word list for Burushaski or any North Caucasian language, I will be happy to perform the Oswalt shift test myself; equally, I will be happy to provide a Swadesh list for Basque to anyone who wants one.)

This failure to recognize the importance of chance resemblances defaces all the work I criticize in my paper: the authors of that work appear to be unaware of the most elementary principles of probability. A magnificent example is provided by Ruhlen, who notes that the Basque and Burushaski words for 'willow' are noticeably similar, and then asks this: "What is the probability that the word for 'willow tree' would consist of two sibilants and another consonant by accident?" The easiest way to reply to this is with another question: What is the probability that the Hawaiian and ancient Greek words for 'honey' will be identical? As it happens, the words are *meli* and *meli*, respectively.

Ruhlen is asking the wrong question. Once we have stumbled across a striking coincidence, gasping in amazement at the *a priori* improbability of that coincidence is a waste of time, and declaring that the coincidence must mean something because it is so striking is a bigger waste of time. The correct question to ask is this: What is the *a priori* probability that *some* Basque word will turn out to be strikingly similar in form and meaning to *some* Burushaski word? This is the kind of question which Ringe has addressed, and the answer is that, providing the two languages have tolerably similar phonotactics, that probability is very high indeed. The coincidence of Basque *sahats* and Burushaski *šAsk,* therefore, is no more interesting than the case of Hawaiian *meli* and Greek *meli*, which after all is the more impressive matchup of the two.

Oh, before someone asks, the Hawaiian-Greek matchups don't stop there. We have Hawaiian *aeto* 'eagle' and Greek *aetos* 'eagle', Hawaiian *kia* 'pillar' and Greek *kion* 'pillar', Hawaiian *mele* 'sing' and Greek *melos* 'melody', Hawaiian *lahui* 'people' and Greek *laos* 'people', Hawaiian *hiki* 'come' and Greek *hikano* 'arrive', and a number of others. Maybe somebody ought to look into this...

10. Systematic correspondences. In my paper, I complain that the comparisons with Basque utterly fail to reveal any systematic correspondences, and I document this charge. The response of the several readers has been quite varied. Pulleyblank is sympathetic: he too would prefer to see "*patterns* of sound correspondence...that would justify the assumption that it is right to ignore the *discrepancies*" [his emphasis]. Wescott argues that the proposed relationship is so distant that correspondences are too much to hope for. Ruhlen says nothing about the issue, and instead adopts what I consider a bizarre approach: he takes the Basque and Burushaski words for 'willow', assumes *a priori* that they are related, posits an arbitrary ancestral form for both, and then invents a sequence of unsubstantiated *ad hoc* phonological changes in each language, including deletions and metatheses, which have supposedly applied *just to that particular word,* for no other reason than that he can in this way derive the attested

forms. This is nothing I recognize as historical linguistics: with this kind of methodology we can derive anything from anything.

Bengtson at first takes a much more responsible approach: he agrees that systematic correspondences are necessary, maintains that he has in fact identified some, and points out (quite properly) that it is unreasonable to expect him and his colleagues to have a complete reconstruction of Dene-Caucasian phonology at this early stage. However, as I declare in my paper and in section 3 above, what Bengtson describes as "correspondences" look to me like nothing of the sort; readers will have to decide for themselves whether they are looking at systematic correspondences or at miscellaneous resemblances.

But Bengtson also does something I consider quite extraordinary. In responding to my objections, he says this: "Most insidiously of all, [Trask] throws in examples that I did not use in my phonology paper". I am dumbfounded to be told that, in evaluating the degree of success of Bengtson's proposed phonological correspondences, I am not allowed to cite any of his comparisons except the ones he has expressly selected to display "in his phonology paper". So, let's see if I've got this right: Bengtson is allowed to offer comparisons that don't exhibit correspondences, and he's allowed to count them as evidence for his hypothesis, but I'm not allowed to criticize them on phonological grounds, because he hasn't licensed them for that purpose. This is the man who complains that his enterprise has found few supporters.

11. Methodology. Like most (not all) of the long-range work I have seen, the attempts at linking Basque to a putative "Dene-Caucasian" family are characterized by a methodology which, in common with most historical linguists, I consider untenable. The responses by some of the readers, in particular those by Bengtson and Ruhlen, have greatly clarified the nature of that methodology, which is now seen to be based upon a number of *a priori* principles which I utterly reject. Here is a list of some of those principles, all of them extracted from the work of Bengtson and Ruhlen, both here and elsewhere, together with my comments on them.

> 1. Basque absolutely must be discoverably related to some other living languages, and the problem is merely to find which those are.
>
> RLT: Nonsense. There is not the slightest reason to presume in advance that a language must have discoverable relatives, even if we accept (as I do) that all languages are probably ultimately related.
>
> 2. Basque must be part of Dene-Caucasian because more comparisons have been advanced between Basque and Dene-Caucasian languages than between Basque and anything else.
>
> RLT: No. If those comparisons are deeply flawed and thoroughly unconvincing, their number is immaterial. A dozen highly persuasive comparisons would mean a lot; 300 bad ones means nothing. The list of Basque-Dene-Caucasian comparisons is longer than other such lists merely because more time and effort has been devoted to compiling it.
>
> 3. In drawing comparisons, it is not necessary to know anything about Basque: merely extracting items from modern dictionaries is good enough.
>
> RLT: Ridiculous. Nobody would try to get away with this approach in comparing, say, English to other languages. Why should Basque be any different?
>
> 4. If the established prehistory of Basque is incompatible with the comparisons, then that prehistory must be changed to make it fit the comparisons.
>
> RLT: Outrageous. If the established prehistory is incompatible with the comparisons, then it is the comparisons which are wrong.

5. It is fair to dismiss even the most knowledgeable, abundant and meticulous work on Basque as "unsubstantiated speculations" if the results of that work are incompatible with Dene-Caucasian comparisons.

RLT: This isn't just arrogant; it's positively offensive. More than any other feature of the work I am criticizing, this is the one that guarantees that no vasconist will ever be able to take it seriously.

6. For a Basque word, a comparison with Caucasian or Burushaski should generally be preferred to a Basque-internal source or to a source in Latin or Romance.

RLT: Preposterous. This is a denial of one of the most fundamental tenets in historical linguistics. A plausible source close to home should *always* be preferred to a remote comparison.

7. If there is a case that a Basque word should be removed from a comparison, adding more languages and more matchups to the comparison greatly increases the likelihood that the original comparison is valid.

RLT: Nonsense. If it can be shown that the Basque word must be removed from the comparison, that is the end of it, and further matchups are irrelevant.

8. Arbitrary segmentations are perfectly legitimate in obtaining matchups.

RLT: No, arbitrary segmentations are to be regarded with the deepest suspicion.

9. When a word exists in several variant forms, the proper procedure is to choose for comparison that form which allows the most persuasive matchup with the other languages.

RLT: Absolutely not. The proper procedure is to choose the most conservative form of the word, as determined in the light of the totality of what we know about the history of the language.

10. Forms from peripheral dialects are always more conservative than those from central dialects.

RLT: Certainly not. It would be a simple matter to provide a long list of innovations which are confined to Bizkaian or to Zuberoan, the two most peripheral varieties of Basque.

This list sums up very well the vast gulf between the methods of those trying to connect Basque to Dene-Caucasian and what I consider to be a proper way of doing historical linguistics. Modestly comparing himself and his colleagues to Rask, Bopp, and Grimm, Bengtson describes my position as "these poor methods", and is convinced that I would not teach such methods to my students. Well, I'm sorry to disappoint him, but in fact my "poor methods" are about to be enshrined in a textbook of historical linguistics which I am currently writing and which, all being well, should be out sometime in 1996. I guess I've probably lost a sale in Minneapolis, but other linguists who have seen my draft chapters seem to think the book is not too bad. It looks like the long-rangers will have to write their own textbook.

The case for relating Basque to North Caucasian, or to Burushaski, or to Sumerian, is not one jot stronger than the case for relating it to Kartvelian (as attempted by Bouda, Lafon, and Braun), or to Berber (as attempted by Schuchardt and Mukarovsky), or for that matter to Hungarian (as attempted in my spoof demonstration in Trask 1994 and Trask forthcoming c). No one who has ever tried to compare Basque to anything at all has failed to find chance resemblances to words in the other language(s) chosen, but so what? This is what I mean when I assert that the evidence for relating Basque to "Dene-Caucasian" is "precisely zero", a statement to which both Ruhlen and Wescott take exception. An exasperated Michelena, in his 1950 review of Bouda's Basque-Caucasian comparisons, rattles off a string of impressive-looking Basque-Indo-European matchups, and, in his 1968 review, he is confident that his list could easily be extended to hundreds of items by anyone determined to pursue the matter.

In fact, it occurs to me now that English is one of the few languages not so far compared with Basque by some ambitious long-ranger. Why don't we try it? Here's a hurried list of comparisons, complete with the now-traditional arbitrary segmentations, especially of those inconvenient initial vowels in Basque.

Basque	English
a-bartsu 'branched'	*branch*
a-bere 'beast'	*bear*
a-burri-tu 'bore'	*bore*
a-fari 'dinner'	*fare*
a-gurtza 'greeting'	*greet*
altz 'alder'	*ald-er*
ama 'mother'	*mama*
anka 'leg'	*ank-le*
a-pal 'lowly'	*pal-try*
auts 'ash, dust'	*ash*
balio 'price, value'	*value*
beha-tu 'look at'	*behold*
bel-tz 'black'	*bl-ack*
bero 'hot'	*bur-n*
bide 'road, way"	*path*
bi-hotz 'heart'	*heart*
burki 'birch'	*birch*
buru 'head'	*brow*
burumuin 'brain'	*brain*
busti 'moist'	*moist*
e-der 'lovely'	*dear*
e-duk-i 'have, keep'	*take*
egi 'border'	*edge*
e-gos-i 'cook'	*cuis-ine*
e-hun 'hundred'	*hund-red*
e-karri 'bring'	*carry*
e-lortza 'thorn bush'	*larch*
e-rbi 'hare'	*rabbi-t*
eska-tu 'ask for'	*ask*
gari 'wheat'	*grai-n*
gibel 'liver'	*giblet*
gris 'gray'	*gray*
gurd-i 'cart'	*cart*
hor 'there'	*here*
i-bar 'valley'	*vale, vall-ey*
i-bai 'river'	*bay*
inozo 'naïve'	*innoc-ent*
i-raul-i 'turn over'	*roll*
i-sil 'silent'	*sil-ent*
iz-otz 'ice' (*hotz* 'cold')	*ice*
j-oan 'go'	*wend*
krako 'hook'	*crook*
mailu 'hammer'	*mall-et*

ma-kur 'curved'	*cur-ve*
multzo 'abundance'	*much*
murmur 'murmur	*murmur*
musu 'snout'	*muzz-le*
neba 'brother'	*nephew*
o-bira-tu 'bury'	*bury*
or-du 'hour'	*hour*
otx 'ouch'	*ouch*
o-ztop-o 'stumble'	*step*
pinu 'pine'	*pine*
pitxer 'jug'	*pitcher*
pix 'urine'	*piss*
put-zu 'well, pit'	*pit*
sa-bel 'belly'	*bell-y*
sei 'six'	*six*
sino 'gesture'	*sign*
sits 'dung'	*shit*
trat-u 'deal, pact'	*treat-y*
u-bel 'weal, welt'	*weal*
u-sain 'odor'	*scent*
zeken 'miserly'	*skin-flint*
zilar 'silver'	*silver*
zintzo 'honest'	*sincere*
zinu 'sign'	*sign*
ziur 'sure'	*sure*
zola 'sole' (of foot)	*sole*

Isn't this fun? And this is only the result of a few minutes' work. What might I achieve with several years of devoted searching for cognates?

Of course, no one would take this seriously, but that's only because everybody knows something about the history of English. If an enthusiastic long-ranger were to protest that such important findings required the known history of English to be radically revised in the light of his data, no one would pay any attention.

So how is the work under review here different? The long-rangers have proposed a large number of comparisons involving Basque and other languages, and concluded that they have identified clear evidence of an ancient genetic link. I have protested at length that the proposed comparisons are utterly incompatible with the established history of Basque. In this I have been supported by every one of the readers who knows anything about Basque, and now even by Bengtson himself. And yet the response of the comparativists is not the expected "Oops!" Instead, they assert hotly that the established history of Basque must be wrong, and that it should be scrapped at once and replaced by something which fits their comparisons better. Apparently we are to believe that Basque, being isolated and little-known to most linguists, is fair game for such shenanigans. In my view, so long as long-rangers continue to adopt this sort of approach, they will continue to find themselves excluded from the linguistic mainstream.

12. Closing remarks. First, I would like to express my gratitude to the editors of *Mother Tongue* for allowing me this extended forum, and for devoting such an extraordinary amount of time and effort to the project. Second, I'd like to express my sincere thanks to all the commentators for taking the time and trouble to read through and reply to an exceptionally long paper. It's been particular fun locking horns with my critics, but I've benefited from all the responses. Finally, I wish the proponents of the Dene-Caucasian hypothesis the very best of luck with their program. I'll be delighted if they can eventually make a convincing case, and I'll be over the moon (as the Brits say) if they can assemble persuasive evidence that Basque must be related to something else. But I regret that I cannot accept as convincing the case which has been made so far.

Additional References

Anderson, James. 1988. *Ancient Languages of the Hispanic Peninsula.* Lanham, MD: University Press of America.

Blust, Robert. 1991. 'Patterns of sound change in the Austronesian languages'. In Philip Baldi (ed.), *Patterns of Change, Change of Patterns: Linguistic Change and Reconstruction Methodology,* pp. 129-165. Berlin: Mouton de Gruyter.

Braun, Jan. 1981. 'Euscaro-Caucasica'. In [no editor] *IKER-1: Euskalarien Nazioarteko Jardunaldiak [Proceedings of the International Congress of Vasconists],* pp. 213-219. Bilbo: Euskaltzaindia.

——— 1984. 'Africadas de las lenguas Kartvélicas y sus correspondencias regulares en el vasco: observaciones adicionales'. In José L. Melena (ed.), *Symbolae Ludovico Mitxelena Septuagenario Oblatae,* vol. II, pp. 875-879. Vitoria-Gasteiz: Instituto de Ciencias de la Antigüedad.

——— 1994. 'Nombres de los insectos en las lenguas kartvélicas y en el vasco'. In Jean-Baptiste Orpustan (ed.), *La langue basque parmi les autres: influences, comparaisons,* pp. 77-85. St. Etienne-de-Baigorri: Editions Izpegi.

Fox, Anthony. 1995. *Linguistic Reconstruction: An Introduction to Theory and Method.* Oxford: Oxford University Press.

Gómez, Ricardo and Koldo Sainz. 1995. 'On the origin of the finite forms of the Basque verb'. In Hualde *et al* (1995), pp. 235-273.

Guiter, Henri. 1968. 'La langue des Pictes'. *Boletín de la Real Sociedad Vascongada de los Amigos del País* 24: 281-321.

Hock, Hans Henrich. 1986. *Principles of Historical Linguistics.* Berlin: Mouton de Gruyter.

Kintana, Xabier. 1981. 'Euskal eta Kartveliar izen batzuren kidetasunaz'. ['On the resemblance between some Basque and Kartvelian nouns.'] In [no editor] *IKER-I: Euskalarien Nazioarteko Jardunaldiak [Proceedings of the International Congress of Vasconists],* pp. 261-267. Bilbo: Euskaltzaindia.

Lakarra, Joseba A. 1995. 'Reconstructing the Pre-Proto-Basque root'. In Hualde *et al* (1995), pp. 189-206.

Maddieson, Ian. 1984. *Patterns of Sounds.* Cambridge: Cambridge University Press.

Michelena, Luis. 1957-58. 'A propos de l'accent basque'. *Bulletin de la Société Linguistique de Paris* LIII: 204-233. Reprinted in Michelena (1988), vol. I, pp. 220-239.

——— 1958. 'Notes sur les éléments romans du basque'. *Via Domitia* V:5-14. Reprinted in Michelena (1988), vol. I, pp. 116-120.

Mistral, Frederic. 1968. *Tresor dóu Felibrige, ou, Dictionnaire Provençal-Français,* 2 vols. Ramoun Berenguié.

Oregi Aranburu, Josu. 1974. 'Euskal-aditzaz zenbait gogoeta' ['Some thoughts on the Basque verb']. *Fontes Linguae Vasconum* 17: 265-283.

Oswalt, Robert L. 1991. 'A method for assessing distant linguistic relationships'. In Lamb and Mitchell (1991), pp. 389-404.

Penny, Ralph. 1991. *A History of the Spanish Language.* Cambridge: Cambridge University Press.

Rijk, Rudolf P. G. de. 1995. '"Nunc" vasconice'. In Hualde *et al* (1995), pp. 295-311.

Ringe, Donald A., Jr. 1992. *On Calculating the Factor of Chance in Language Comparison. Transactions of the American Philosophical Society* 82(1).

Trask, R. L. forthcoming *d.* 'On finding relatives for Basque: what vasconists would like to see'. To appear 1996 in *Acta Linguistica Hafniensia.*

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

POSTSCRIPT: RESPONSE TO JACOBSEN

For the reason he explains, Jacobsen's extremely interesting comments reached me after I had already sent in my response to the other readers' comments; because of exigencies of time and space, I am unable to do them anything like full justice here. I shall therefore single out what I consider to be the most interesting remarks. Jacobsen notes and corrects a few minor errors and omissions in my paper; I am happy to accept these corrections.

Phonology. In my main response, I am severely critical of the several comparativists for their reliance on modern-day Basque forms and their refusal to take into account the established reconstruction of the pre-Basque of some 2000 years ago. I now find that Jacobsen is proposing a phonological system significantly different from that reconstructed by Michelena, for an earlier stage than that of Michelena's reconstruction, and that he is therefore also proposing a large number of developments (in pre-Roman times), particularly affecting initial plosives. Naturally, he does not defend his views in any detail here, and I shall therefore await a fuller exposition with interest. For the moment, I note only two points. First, his proposals appear to offer no support to the comparativists whose work I am criticizing. Second, if Jacobsen proves to be successful in making a case for his posited phonological system, the result will be that the ancestral forms of Basque words will look even more different from the modern forms than they already do, undermining still further comparisons based exclusively upon modern-day forms.

I find myself unable to accept Jacobsen's suggestion that pre-Basque might have had a phoneme /m/ after all. I do, however, agree with Jacobsen in concluding that [b] and [m] did not constitute separate phonemes. Recall that there is abundant evidence that [m] was an allophone of /b/ in certain positions, notably word-initially: as I point out in my main response, virtually every native Basque word in *m-* contains, or formerly contained, a following /n/.

I do not understand the appeal to the meaningless prefix *m(a)-* (better *ma-*, in my view). Its very meaninglessness, as in *madari* ~ *udari* ~ *udare* 'pear' and *magal* ~ *hegal* 'wing', points strongly to a purely "expressive" function: in several places, Michelena points out that the rarity of /m/ led to its being favored in expressive formations, as it still is today. Moreover, in the very passage cited by Jacobsen, Michelena notes that the words with *m-* often exhibit regional variations in form which are unparalleled by ordinary lexical items, and that they often have anomalous phonological characteristics like medial *-lk-*, again underlining their purely "expressive" status. And I do not object to the putative prefix **b(e)-* on the ground that it is "meaningless": I object on the ground that there is no evidence for its existence.

As for the words in *m-*, like *moko*, *moto*, and *mutur*, my main response already points out that such words absolutely do not look like native Basque lexical items: there is probably not a single indisputably ancient native word of the form **mVPV-*, where *P* is a voiceless plosive, and words of this form therefore cannot safely be taken as ancient lexical items: they must be either loan words or expressive formations. The words with medial *-m-* I have already dealt with.

I am sympathetic to the idea that initial **d-* was systematically converted to something else before the Roman period, but I would question the conclusion that it changed to *n-*, which is itself a rare initial in native words. Nevertheless, I recognize the value of Jacobsen's proposal in accounting for the first-person alternations, though I wonder what he makes of *ene* 'my'.

Jacobsen suggests that medial *-p-* was probably absent in pre-Basque; I agree that this is possible, but I note cases like *lepo* 'neck', *ipurdi* 'buttocks', *epel* 'lukewarm', and *eper* 'partridge', which require an explanation.

Jacobsen agrees with me that native Basque words do not begin with voiceless plosives. This is a crucial point, and one absolutely fundamental to my case, since the linguists whose work I am criticizing explicitly make the opposite assumption.

On the vexed question of the origin of the aspiration, Jacobsen declares that I differ from Michelena, who "clearly assumed that a phoneme **h* was found in Proto-Basque". I must take exception: this is not Michelena's position, and I do not differ from Michelena. It is true that, on occasion (e.g., 1977:205), Michelena idiosyncratically represents the pre-Basque aspiration as "/h/", as though it were a phoneme. But his reconstructed phoneme system (1977: 374 and elsewhere) does *not* include an /h/, and his discussion of the aspiration in Chapter 21 of his book centers on the following claim: "la aspiración está condicionada por la posición del acento antiguo"

/"the aspiration is conditioned by the position of the ancient accent"/ (p. 418). Michelena and I agree entirely: pre-Basque had a phonetic aspiration, but this aspiration was not phonemic; it was of suprasegmental origin, and it did not continue an earlier segment.

Jacobsen imputes to me the view that the occurrence of the aspiration in loan words is "hopelessly perverse". It was never my intention to defend such a position, but I guess I wasn't very clear. What I mean is the following. First, the position of the aspiration is usually predictable, which is incompatible with the insistence of Bengtson and others that it continues a segmental phoneme. Second, as a consequence, it occurs in loan words in positions in which it cannot possibly continue an original segment. Third, however, its position is sometimes quite variable even within a single word, so that a segmental origin for it would require some startling metatheses. All of this is intended to reinforce Michelena's (and my) conclusion: the Basque aspiration does not continue an earlier segment.

Jacobsen himself goes on to demonstrate rather effectively that the long-rangers have invoked the Basque aspiration, quite illegitimately, to match *almost any* consonant in the other languages, and that, even so, they have very frequently matched it with zero in the other languages. It is one thing to speak of "mergers" or of "loss of conditioning factors"; it is quite another to treat a single segment as a magical toy which can fill any slot at all and which can also quietly disappear when its presence is embarrassing.

I am puzzled by Jacobsen's flat denial of my posited **onhore* 'honor' as the precursor of modern *ohore* (and variants). Such a form would be exactly parallel to the treatment of other words containing resonants, such as *alhaba* 'daughter', *belhar* 'grass', *erhi* 'finger', and *urrhe* 'gold', not excluding loan words like *sorho* 'field'. The loan *anhoa* 'provisions < ANNONA exhibits such behavior quite overtly, and Jacobsen's attempt to explain this by a metathesis of an earlier **anoha* requires an aspiration in the third syllable, something which (apart from compounds) is unattested and apparently impossible. But nothing of consequence appears to hang on this.

Individual comparisons. In the majority of cases, Jacobsen's observations support my rejection of the proffered comparisons; in several instances he draws attention to further difficulties with the comparisons which I overlooked. I have only a few remarks to make here.

Item [55.] Jacobsen considers that *beko* and *moko* have distinct origins; on reflection, I suspect that he is right.

Item [80.] The phonologically anomalous verbal prefix *d-* is certainly a puzzle. But I cannot accept Jacobsen's proposed pronoun **do:* as stressed in my main reply, *d-* is a tense marker, not a person marker.

Item [90.] Jacobsen concludes, correctly I think, that Basque *eme* 'sweet' is nothing but a specialization of the loan word meaning 'female'.

Item [100.] Like me, Jacobsen sees Basque *etxola* 'hut' as a compound of *etxe* 'house' and *ola* 'hut'. This view is endorsed in several places by Michelena; elsewhere, however, Michelena suggests that *etxola* in fact results either from a cross between *etxe* and the loan word *txa(b)ola* 'hut' or from a compound of these elements (Michelena 1970a: 42 fn.).

Item [243.] Jacobsen's suggested etymology for Basque *pataxa* 'bottle' is greatly superior to mine, and I am happy to accept it.

Item [245.] Jacobsen points out that Basque *pintza ~ mintz(a)* 'membrane' has no convincing Latin source. The word is nonetheless well attested in Romance, and I consider a Romance loan into Basque to be far more likely than the reverse, on the ground of precedent.

Segmentation and affixes. Jacobsen follows me in condemning the arbitrary segmentations of the long-rangers, pointing out correctly that projecting morpheme boundaries from one language to another constitutes circular reasoning.

Jacobsen attempts to put some rigorous content into the old proposal of a "body-part prefix" in Basque. He considers that his revised version is "a little more favorable" to the comparisons adduced, but that this prefix "has nothing to do with Caucasian prefixes". This is cold comfort for the comparativists.

I cannot accept the suggestion that **-el* might constitute a suffix in the names of internal organs; as my main reply makes clear, *gibel* 'liver' is clearly to be segmented as **gi-bel* and *sabel* 'stomach' is probably to be segmented as **sa-bel.*

Evaluation and conclusions. Jacobsen is expressly sympathetic about the difficulty of identifying systematic phonological correspondences at the kind of time depth under consideration here, but he nonetheless makes a few pointed remarks about the brevity of the morphs which are adduced as matches.

In my paper, I concentrate on two weaknesses of the proposed comparisons: the lack of phonological correspondences and the arbitrary (and often clearly wrong) segmentations. Jacobsen here turns to a third difficulty, one which I did not discuss: the lack of constraints upon the semantics. Since the long-rangers make little or no attempt to control phonology or morphology, it might seem reasonable at least to impose some fairly strict controls upon the semantics, but Jacobsen finds only what he calls "semantic discontinuity": the matching of words which have entirely distinct meanings but which happen to fall within a broad semantic domain.

Naturally, he objects to this. Comparing words for 'daughter' and 'sister', or words for 'old woman', 'female', and 'paternal aunt' must, of course, have precisely the consequence that Jacobsen points to: the multiplication of chance matches in form. This, of course, is another aspect of that methodology which I condemn so strongly in my main reply: if you choose some languages for comparison, assuming in advance that they are probably related, and if you then proceed to collect possible confirming instances, with no controls and no scrutiny of your results, then of course you will find bunches of matchups.

Jacobsen goes on to make a further point. In nearly half of the comparisons on offer, a Basque word is compared with a word in only one of the other languages or families considered (I have, of course, omitted the Sino-Tibetan and Na-Dene data, but adding these data would not change the overall picture). Moreover, in comparisons with Caucasian, the great majority involve comparisons with only one of the three recognized North Caucasian groupings, and quite a few involve comparisons with only a single language.

Jacobsen's conclusion, in his final sentence, speaks for itself.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

CANAANITE VOCABULARY IN BENGALI AND IN SOME OTHER IE DIALECTS OF INDIA

Liny Srinivasan and Cyrus Gordon

[Editor's Note: Liny Srinivasan has the doctorate from University of Pittsburgh in Geography of Religion. An American citizen, originally of India, and a native speaker of Bengali, she has concentrated on Indian religion and history, receiving a Fulbright Scholarship for field work in India. Her current address is 224 Clinton Avenue, Eatontown, New Jersey, USA 07724. Cyrus Gordon is a well-known Semiticist, authority on Ugaritic and Eblaite, the Minoan as Semitic hypothesis, and Professor emeritus at both Brandeis and New York University. His address is 130 Dean Road, Brookline, MA, USA 02146-4212.]

Liny Srinivasan has observed that "Desi" words (which are all of non-IE origin) in the Indic languages of India (notably Bengali) are usually of Canaanite derivation. She was led to this conclusion by literary references to Kanyani-vasina "dwellers of Kanyani". She realized that the collaboration of a Semiticist was necessary to separate the wheat from the chaff. On the advice of Dr. Nelly Segal of the Rutgers University Library and of Monmouth College, she was directed to her (Nelly Segal's) former teacher, Cyrus Gordon, a Semiticist specializing in Canaanite, including Ugaritic and Hebrew. On reading Srinivasan's unpublished typescript, it was clear to him that her important discovery was correct, though for proper publication, it needed to be sanitized in accordance with phonetic, semantic, and other linguistic criteria.

It is not enough to assert that Kanyan is "Canaan" just because they sound somewhat similar. The question is: Are there other Desi words with *-y-* corresponding to the laryngeal ʕ in Canaanite? Srinivasan correctly identified D[esi] *beyara* with H[ebrew] *ba-a-ra*, because they both mean the same thing, 'silly, foolish'. The H form is actually *baʿar-*. When Cyrus Gordon saw that here too, Canaanite ʕ comes into D as *y*, he no longer hesitated to accept Srinivasan's identification of Kanyan with H *Kenaʿan* 'Canaan'.

The preceding paragraph illustrates the essential nature of the collaboration between Srinivasan and Cyrus Gordon in establishing the Canaanite origin of Desi words.

The contacts between the Semitic Near East and India are spread over millennia (down to the present) and have therefore left effects in many layers on the two regions. The first seminal contact that can be dated with firm linguistic evidence goes back to about 2000 BCE, as will be indicated below. But the main thrust of this article is to present enough D words of Canaanite origin to show that the evidence is massive, and by no means tenuous.

The phonetics of the Devanagari script is not compatible with Semitic phonology as represented in Near East types of writing. The presence of, or even the existence of, aspiration, laryngeals, emphatics etc. make it impossible to square the two phonological systems. Any comprehensible one-to-one correspondence is out of the question.

D[esi] *ṣaṛ* 'bull, ox' is derived from Semitic. Cf. H[ebrew] *šor*, Arabic *ṯawr*, Ugaritic (written consonantally) *ṯr*, etc. The Aramaic form *tor* < *tawr*, was borrowed into Greek as *taur-os* and into Latin as *taur-us* 'bull, ox'. The borrowing went hand in hand with the westward spread of ritual bull-grappling, bull-vaulting ,and bull-fighting. In Spain, the bull-fighting season begins on Easter Sunday so that the ancient religion is coupled with the current official religion. The very terminology of 'toreador' and 'toro' (bull) still remind us of the ancient Semitic origins.

D *sita* 'winter' < Semitic: Arabic *šitaʾ*, H *stav* 'winter, autumn, the rainy season'.

D *gol* 'round, globular' < H *ʿagol* 'round, globular'. Srinivasan points out examples where the initial syllable of the Semitic original is dropped in D; e.g., D *gali* 'a narrow way, path' < H *maʿgal* 'path, way, track'. Similarly, D *khal* 'canal, brook' < H *naḥal* 'brook, stream (which flows only in the rainy season)'. [There are four *h*'s in Egypto-Semitic: *h*, *ḥ*, *ḫ*, and *ḫ*. The last [ḫ] is distinguished graphically only in Egyptian.] Other examples of

dropping the (original Semitic) initial syllable in D words are available among the following:

D *mita* 'friend'; cf. H *ʿamit* 'friend, neighbor'.

D *cheli* 'purple red ritual garment'; H *tekel-et* (garment) dyed blue or purple'. (The *-et* is an ending of feminine singular nouns.) In H, for over a thousand years, an undoubled *k* (i.e., not *kk*), when immediately preceded by a vowel, is spirantized (and pronounced [x]). Such technical details must be borne in mind to understand Srinivasan's rendition of D words in Latin letters. Her transliterations and translations of D words have not been modified in this article.

D *fola* 'swelling' < H *ʿofel* 'swelling' (applied to places where the terrain is in the shape of a swelling, like the Ophel of Jerusalem).

D *tham* < H meaning 'to stop, to be perfected, finished' is also used to negativize verbs. The Egypto-Semitic root *tmm* 'perfect' (noun and adjective) is also used to negativize verbs. E.g., Coptic *ptemkalaunan* 'not letting anybody-or-anything at us = inaccessability'. It is composed of *p* 'the' + *tem* 'not' + *ka* 'letting' + *lau* 'anybody/anything' + *nan* 'to/at us'.

D *hoi/haya* 'is, are' < H *haya* 'to be, exist'; e.g., *haya* 'he was, *hayu* they were'.

D *bagan* 'garden'; cf. H *ba-gan* 'in the garden'; composed of *ba* 'in' + *gan* 'garden'.

D *bas* 'cloth, garment'; H *lbš* 'to put on clothing, to dress'; *lebûš* 'attire, dress'. As noted for post-vocalic *k*, the consonants *b g d k p t* (unless doubled) are all spirantized when preceded directly by a vowel. Accordingly, this word for 'attire' is pronounced *levûš*. However, in the sequel, such linguistic finesse will generally be omitted because (for the non-Hebraist), it tends to obscure rather than clarify the subject in hand (and for the Hebraist, it is unnecessary). Vocalic length is generally not indicated for the same reasons.

D *thoka* < H *taqaʿ* both meaning 'to drive in a nail'.

More D words derived from H will be listed further on.

Srinivasan has stated that a people called Abhira have idioms derived from H phrases. This induced Cyrus Gordon to investigate whether "Abhir-" was related to the name of the "Hebrews", which is *ʿibri* in Hebrew, *ḫa-bi-ru* in Akkadian, and *ʿpr* [ʕapir] in Egyptian (where the plural *ʿprm* was vocalized [ʕapirûma]). The discrepancies indicate that the word is of non-Semitic origin. In Sanskrit, there is unaspirated *b* and aspirated *bh*, as well as unaspirated *p* and aspirated *ph*. In Semitic (as in English), *b* is regularly unaspirated, and *p* is regularly aspirated. There is no way to write *bh* in Near East scripts. The Hebrews chose to represent the sonant quality, which meant giving up the aspiration. The Egyptians chose to indicate the aspiration, which meant giving up the sonant quality of the *bh*.

It is interesting to note Indic *bhang-*, which designates either 'hemp fibre' or 'the narcotic made from the leaves and tops of the same plant.' *Bhang-* occurs in Hebrew (Ezekiel 27:17), where it is spelled consonantally *p-n-g* because there is no letter for the aspirated *bh* in the Hebrew alphabet. The roles of both the Abhira and of the Near East Hebrews *ʿprm* are varied. Abr(ah)am the Hebrew (Genesis 14:12) is the progenitor of God's special People). (Similarly, in the Mahabharata, the sixth son of Kalaka, a wise and great Asura, was on earth the great royal seer Abhiru.) In Egypt, the Hebrews were slaves. In the Nuzi tablets (during the Amarna Age), they often, as individuals, voluntarily enter slavery. At the same time in Palestine, marauding bands of *Ḫa-bi-ru* are wresting Palestine from the Egyptian rulers. This topic, which is of broad interest, will be treated in detail with adequate documentation elsewhere.

At this juncture, we should note that the Hebrews have recorded their contacts with IE speakers. Genesis 23:6 not only records a transaction between Abraham and the Hittite community presided over by Ephron the Hittite around Hebron. It represents Abraham as a prince in the service of the deified Emperor (which is what "you are a prince of god in our midst" means).

Hittite is the first known written IE language. All of Syria-Palestine was called Hittite-land; e.g., for centuries in the Assyrian Annals of the Iron Age (which begins circa 1200 BCE). Later, around 600 BCE, the Prophet Ezekiel (16:3) tells the population of Jerusalem: "Thy father is the Amorite; and thy mother, Hittite". This implies that the two main elements that hybridized to produce the biblical Hebrews, were the West Semitic Amorites and the IE Hittites. This has been foreshadowed for decades in publications such as Cyrus Gordon's "Indo-European and Hebrew Epic" (*Eretz-Israel*, vol. V, 1958, pp. 10-15). Thus the institution of levirate marriage in Genesis 38 is not in keeping with biblical and postbiblical Jewish cases and laws, but has IE factors alien to all Jewish usage. Note that the father-in-law (Judah) fills the role of *levir*, which is permitted only in one known code of the ancient Near East; namely, the Hittite Code (section 193). While the very idea of burning the widow is contrary to anything Jewish, it smacks of *sati* 'suttee': Indic widow-burning. In all the historic stages of Jewish levirate marriage, biblical and postbiblical, the levir takes the place of his dead brother and continues indefinitely to live with the widow as man and wife. But in India, the pair ceases to have sexual relations as soon as the son is produced to continue the line of the brother who died childless. In all of Hebrew literature, it is only in Genesis 38:26 that the man (Judah) never touches the woman (Tamar) after impregnating her with the heir (Claire Gottlieb, *Varieties of Marriage in the Bible: And their Analogues in the Ancient World*, University Microfilms, Ann Arbor MI 48106, 1989, p. 166).

The linkage between the Northwest Semites and the Indo-Europeans can be documented back to circa 2000 BCE. Moreover, it is from the cuneiform tablets of the Near East that we have not only IE Hittite records but also specifically Sanskrit idioms many centuries before the earliest known Sanskrit texts from India. Thus, there is a common background shared by Indic and the Near Eastern civilizations. For all we know, the pre-Indic Sanskrit (alluded to above) may antedate the IE invasions into Tamil India. But before the actual evidence of this far-reaching fact is presented towards the close of this article, the reader should be acquainted with more Semitic Desi words, which are opening a Pandora's Box of discoveries and problems that bid fair to keep a rising generation of linguists on their toes and very busy for a long time.

D *bana* 'to build, construct'; < H *bana* 'he built' (root *bny*); from the same root is D *ben*, *biyan* 'child birth' and H *ben* 'son, child'.

The H root *gll* > D *golala* 'rounded' and *gola* 'round balls'.

The H root *dʔy* 'to fly, soar' (as in *daʔa* 'he flew') > D *dha* 'to run, move swiftly as flying'.

D *tola* 'to draw up water' < H *dly* 'to draw up water/fluid (e.g., *dala* 'he drew water'). From the same root is H *deli* 'bucket' which could well be the source of D *dala* 'a vessel'.

H *dʔg* is the root of words indicating 'care/anxiety/worry'; e.g., the noun *deʔaga* 'care, anxiety, worry' which is the source of D *daga* 'to cause mental anxiety'.

The H root *dḥl* 'to fear, frighten' may well provide the origin of D *dhakal* 'trouble, misery'.

H *ḥalaq* 'smooth' with many applications including the 'smooth talk' associated with flattery and diplomacy. It explains the origin of D *chalak* 'smart, diplomatic'.

H *ḥamor* 'donkey' when applied to people, means 'stupid'. It is so used in modern Hebrew, as is its cognate *ḥimar* 'donkey' in Arabic. Cf. D *chamar* 'a low caste, an abusive term'.

H *ṭippeš* 'fool, foolish (root *ṭpš*) provides the origin of D *dhapas* 'fat, inert, stupid'.

H *krt* is the root 'to cut off'; e.g., *karat* 'he cut off'. Cf. D *karat* 'a saw'.

H *maḥa* 'he erased' (the root is *mḥy*); provides the origin of D *mocha* 'wipe, wash, cleanse, erase'.

H *migdal* (the older form is *magdal*) 'tower' > D *magdal* 'the loftiest / topmost part of a building or tree'.

D *sara* 'to turn aside, to go away' < H *sûr* 'to turn aside, go away' (e.g., *sâr* 'he turned aside, *sâra* 'she turned aside').

H *palîṭ* 'a fugitive, an escapee' (root *plṭ*) provides the origin of D *pala* 'to flee, escape'. The *ṭ* is reflected in the Sanskritized form of the D; to wit, *palatak* 'a fugitive'.

H *pilles* 'to level, straighten out' and *peles* 'scale, balance' provide the origin of D *palla* 'the beam and balance of a scale'.

H *petaḥ* 'the opening of a dwelling (especially of a tent)' (root *pṭḥ* 'to open') > D *phatak* 'main door, gate of a house, fort'. In Old Bengali, *phitoh* means 'to open'.

H *pzr* 'to scatter'; e.g., *pizzer* 'he scattered'. This provides the origin of D *pa-char* 'to disperse, distribute, send goods from one place to another secretly (often stolen goods)'.

H *ṣaḥ* 'pure' (masculine singular), *ṣaḥḥa* (feminine singular); root *ṣḥḥ* provides the origin of D *sach, sachcha* 'pure, genuine, true'.

H *ṣar* (masculine singular), *ṣarra* (feminine singular), *ṣarru* (verbal masculine plural) 'narrow' (root *ṣrr*) > D *saru* 'narrow'.

H *qeʿara* 'dish (= *qeʿar* + feminine singular suffix) > D *karʔai, karʔa* 'a cooking bowl, wok'.

H *qaš* 'straw, stubble' > D *kaŝ* 'a reed grass'.

H *qana* 'he bought' (root *qny*); the infinitive is *qenot* 'to buy, purchase' > D *kena* 'to buy, purchase'.

H *šûṭ* 'to go to and fro' (root *šwṭ*) > D *choot, chuta* 'to run'.

The two above lists of Desi words that are derived from Semitic (and mainly from Canaanite) should suffice. The number of such words could be multiplied here and now, but such "overkill" would not serve any useful purpose in this article. Instead, we now aim at clarifying the importance of the subject as a whole, as far as we can grasp it at this moment. Of one thing we can be sure: There will be much more to say during the coming months and years. This article opens the subject; it does not close it.

The IE Hittites introduced the horse and chariot to the Near East circa 2000 BCE, revolutionizing the art of war and changing the course of history.

Johannes A. H. Potratz (*Die Pferdetrensen des Alten Orient, Analecta Orientalia* 41, 1966, Rome: Pontifical Biblical Institute) provides a thorough treatise on the arts of training chariot horses, of the construction of chariots, and of harnessing in the second millennium BCE, with copious illustrations (plates and drawings).

Annalies Kammenhuber (*Hippologia Hethitica*, 1961; Wiesbaden: Otto Harrassowitz) provides a

masterful, annotated edition of the most important of the hippological texts unearthed in the Imperial archives at the Hittite capital of Hattusas (modern Boghazköy, in Turkey). That text was composed by Kikkuli, the horse-trainer from the land of Mitanni, an important Hurrian kingdom in the bend of the Euphrates. It was independent during the early part of the Amarna Age (circa 1400 BCE) until it was conquered by the Hittites and incorporated into their empire. The text, which is in the Hittite language, opens with the statement: "So says Kikkuli, the horse-trainer of the land of Mitanni."

The Hurrians were an important ethnic factor in the Near East. They excelled in a variety of roles including religion. For example, at Ugarit there was found a tablet with a hymn to the Sun goddess. Although the majority of the population at Ugarit (including the king) were Semites, the hymn is in the Hurrian language. Indeed, the Egyuptians often designated Syria-Palestine as "Ḫurru-land'. At Ugarit, the trilingual vocabularies have Sumerian in the first column, Akkadian in the second column, and Hurrian in the third. It is only in the quadrilingual vocabularies that a fourth column supplies the official language of the realm; namely, Ugaritic. The religious functions of the Hurrians gave them a status anticipating the Brahman caste of India.

Another role of the Hurrians was their leadership in the art of war, including the development of the horse-drawn chariot, as exemplified in the Kikkuli text. The latter, as we shall soon see, has specifically "Indic" features. The Hurrian military elite anticipates the Kshatriya caste of India.

While the Hittite language as used in the Kikkuli text is IE, it contains terminology that Kammenhuber calls "Aryan" but which other scholars would call the Indo-Iranian branch of IE. The terminology in question designates the turn in the course around which the horse drew the chariot.

The "turn" is called *wartanna* from PIE **wert-*, Aryan **vart-* 'to turn'. The turns that are numbered in the text are "turn one", "turn three", "turn five", "turn seven", and "turn nine". Thus we have the Aryan odd numbers from one to nine. These numbers, which are not in the IE Hittite language but definitely in Aryan (very close to Sanskrit), are as follows:

1. *aikavartanna* "the one-turn". Sanskrit has *eka* '1' where the Kikkuli has *aika.*

3. The "three-turn" is spelled *ti-e-ra-wa-ar-ta-na = tervartanna*, where the Indologists would expect to find **tri-vartanna.* (Such dialectal differences must be taken in our stride, when we deal with new sources that open new horizons).

5. The "five-turn" is spelled *pa-an-za-wa-ar-ta-an-na*, which is normalized as *panca-vartanna*, with the numeral '5' pronounced like Sanskrit *pañc(a)* '5'.

7. The "seven-turn" is spelled *sa-at-ta-wa-ar-ta-na = satta-vartanna.* Sanskrit has *sapta* '7' which is the original form; cf. Greek *hepta*, Latin *septem.* Kikkuli *satta* with **p* assimilated to the following *t* is a secondary development.

9. The "nine-turn" is spelled *na-a-wa-ar-ta-an-na = navartanna*, where we might have expected **nava-vartanna.* Perhaps the Kikkuli form has undergone haplology with *va* for *va-va.*

In any event, we have in the language of the numbered turns of the Kikkuli text, "Sanskrit" in the Near East, many centuries before Sanskrit was written in India. I (Cyrus Gordon) could say much more at this time but the above should suffice to show the importance of Srinivasan's discovery that the Desi words in Bengali are predominantly of Canaanite origin. It was that discovery which touched off the chain reaction.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

THE CONCEPT OF PROOF IN GENETIC LINGUISTICS

Joseph H. Greenberg
Stanford University

1.0. The Notion of Proving Relationship

Virtually all historical linguists share a common notion which may be stated in something like the following terms. One starts with a hypothesis of relationship between two or more languages or groups of languages.

If groups of languages are involved, these in turn are conceived of as a number of languages which already have been proven to be related, e.g., the Indo-European family. Taking as examples binary hypotheses, which are the most frequent in the literature, we may illustrate the three possibilities by actual examples. An instance in which we deal with two single languages is the Japanese-Korean hypothesis, of a single with a group of languages, the Eskimo-Indo-European hypothesis, and of one group with another group the Indo-European-Uralic hypothesis. Note that in formulating it this way, what we call a group of distinct languages is basically a relative matter. Virtually every language has internal dialect divisions, and the point at which we talk of a group or family as distinct from a single language is to some extent arbitrary. Moreover, some of these hypotheses as usually stated take some single language as part of the hypothesis because it is more important or better known and disregard the group of which it is a member. Thus, Uhlenbeck, who championed the Eskimo-Indo-European hypothesis, treated Eskimo *as though* it were a single language, using West Greenlandic, the best studied form, as representative of Eskimo as a whole. In fact, Eskimo contains at least two units, Inuit and Yuit, which certainly deserve to be called separate languages. In addition, Eskimo is universally recognized as having Aleut as its nearest relative in a language family called Eskaleut. It is noteworthy in such typically binary hypotheses, that less important languages are generally disregarded. Thus, Aleut is almost never compared with anything else while the more prominent Eskimo is. So also for the purpose of Japanese-Korean comparison, the dialect group found in the Okinawa Islands and which certainly deserves the status of a separate language is generally disregarded in comparisons of Japanese with other languages or language groups.

2.1. Probability versus Demonstration

When hypotheses such as those mentioned in the previous section are advanced, what is sought is "proof", and indeed phrases like "to demonstrate" the relationship of A to B, a terminology which is, of course, ultimately taken from geometry, are frequently used. What is sought is certainty of the kind attainable in mathematics and logic. I believe there is at least a dim realization that in all empirical sciences, as against logic and mathematics, in which truths flow infallibly and tautologically from definitions, all that we *can* get are results so close to certainty that for all practical purposes we can consider them true, that is, a hypothesis which is overwhelmingly better than any other in accounting for the facts. Such hypotheses have a further important characteristic, which we may call fruitfulness. That is, they can be built upon and lead to further discoveries and explanations of hitherto puzzling phenomena. In the search for infallibility, certain criteria have been advanced which some linguists think will bring them results which are tantamount to certainty. The two most popular are the existence of sound correspondences and the reconstruction of a proto-language from which the languages hypothesized to be related can be derived. Both of these will be considered later. For the moment, it is to be noted that the two approaches involve somewhat different metaphors from different non-linguistic fields. Sound correspondences are often called laws, or more exactly, the sound changes that give rise to the correspondences, are the so-called "sound laws". Here the analogue is to the laws of nature, which, like sound changes, brook no exceptions. They thus vicariously share the certainty and prestige of such laws as the law of gravitation, one which is of course an empirically founded one and which in its Newtonian form was ultimately shown to be a special case of the more general principle of relativity. The

metaphor of reconstruction is more like that of logic. The reconstructed forms are like postulates, and the changes likes laws of deduction. Thus, the forms to be explained are derived by a procedure which has some analogy to that of deductive logic and in this way participates in its prestige and certainty.

As so often, Aristotle hit the nail on the head. In the *Nicomachean Ethics* (I.3.4.), he states: "For it is the mark of an educated mind to expect that amount of evidence in each area which the nature of the particular subject admits. It is equally unreasonable to accept merely probable conclusions from a mathematician and to demand strict demonstration from an orator." In my mind, linguistics is somewhere in the middle on such a tacit scale, the most humanistic of the sciences and the most scientific of the humanities.

2.2. Analytic and Synthetic Truth

What we have been sketching is, of course, akin to the famous Kantian distinction between analytic propositions, true by definition and synthetic, drawn from experience and therefore only subject to lesser or greater degrees of confirmation. All the twisting and turning in the world cannot make reconstruction or sound laws give the complete certainty attaching to the truths of the disciplines of logic and mathematics.

It is crucial to note that the view of genetic linguistics as concerned with proving relationships and the notion of certain procedures as providing such proof are shared by the "conservatives" who believe that there are a large number of independent, or at least not provably related families in the world, and those who undertake long-range comparisons like the Nostraticists. It is just that the "proofs" of the latter are not accepted by the former as adequate. The Nostraticists, in particular, tend to belong to the camp of the reconstructionists, that is, those who believe that a reconstruction of an ancestral language proves the relationship of the descendant languages.

2.3. The Relativity of Proof

The existence of numerous points of disputes concerning language classification, so well-known to all historical linguists, has just been alluded to. What it shows is that comparative linguists do not have a well-formulated and generally agreed on notion of what in fact constitutes proof of a hypothesis of relationship.

If one extends one's view backward from contemporary linguistics to earlier periods, the differing notions of what proof is becomes even more striking. Consider the following statements from an earlier but not remote period regarding the Finno-Ugrian family, actually along with Samoyed, a subgroup of Uralic, a family universally accepted.

The first comes from Kai Donner, an eminent Finno-Ugricist and the founder of the journal *Finnisch-Ugrische Forschungen*,

> Through J. Sajnovics (*Demonstratio* 1770) i.e. *Demonstratio idioma ungarorum et lapponum idem esse* and S. Gyarmathi (*Affinitas* 1799) i.e. *Affinitas linguae ungaricae cum linguis Fennicae originis grammatice demonstrata*, Strahlenberg's well-founded assertion regarding the Finno-Ugrian (Uighur) group was proven once and for all. (1901:129)

The reference here is to a work published by Strahlenberg in 1730.

Later than Donner, a well-known Finno-Ugricist Ravila (1935:21) stated,

> In 1770 the Hungarian Sajnovics published his famous *Demonstratio* in which, using quite modern methods he proved the relationship of Hungarian to Lapp. By this and by the *Affinitas* of Sámuel Gyarmathi which appeared in 1799 the Finnish-Hungarian relationship was regarded as finally established.

That this view is widely held by Finno-Ugricists up to the present is shown further by a review of an English translation of Gyarmathi written by Z. McRobbie (1986:159), in which Gyarmathi's achievements are summed up in the following terms:

> Sámuel Gyarmathi was the first scholar to analyse all of the Finno-Ugrian languages on a broad basis by emphasizing systematic lexical and morphological comparisons. He was able to define degrees of linguistic affinity pointing out that Vogul and Ostyak are the closest relatives to the Hungarian language. And although Gyarmathi did not utilize regularities in sound correspondences he was nevertheless able to establish a number of Finno-Ugrian etymologies still valid today.

Gyarmathi's work of course preceded the major work of Bopp in 1816, generally viewed as the beginning of comparative linguistics in relation to Indo-European. It preceded the first reconstruction of PIE by Schleicher in the 1850s and of course the first formulations by the Neogrammarians of the notion of exceptionless sound laws in the 1870s. Yet, as we have seen, Gyarmathi's work has been widely viewed by Finno-Ugricists as having proved the relationship of these languages.

At the other extreme, and very recently, Callaghan, in a review of a dictionary by Harvey Pitkin of Wintun, a Penutian language of California remarks,

> There has been a long debate about whether all or part of *none* (italics mine) of the Penutian hypothesis is valid. A preliminary step to the determination of remote relationships is competent reconstruction of proto-languages at a lowest level. (1991:131)

Presumably, Penutian includes here California Penutian. Except for the extinct Costanoan, whose likeness to Miwok is obvious, we have one or more dictionaries and grammars of all the languages Miwok, Yokuts, Wintun, and Maidu. What more evidence do we need to decide the validity of California Penutian at least, which was proposed by Dixon and Kroeber in 1919, more than seventy years before Callaghan's statement? Moreover, Callaghan (1958:193) herself had stated that a consensus of opinion among the four field workers including herself concerned with these languages was that California Penutian does indeed constitute a linguistic stock. In the very same issue of *IJAL*, there was an article by Pitkin and Shipley, two of the field workers alluded to in the preceding statement, which contained over 400 etymologies of California Penutian with the usual apparatus of sound correspondence and reconstructed forms.

I have purposely counterpoised here two extremes, the "prescientific" work of Gyarmathi on Finno-Ugrian, widely accepted as "proof" of the validity of that stock, and the evidence apparently accepted by Callaghan in 1958, but rejected by her in 1991 as inadequate, since she states that possibly none of the Penutian hypothesis is correct.

From this, we see what constitutes proof of relationship varies tremendously, depending on the particular language group involved and the historical period. There is then no consensus as to how much and what sorts of evidence is sufficient to "prove" relationships. What we actually have at present as seen in encyclopedias and general reference works on linguistics is a product of historical accident. Objectively, the California Penutian languages are probably closer to each other than Finnish is to Hungarian. If a stock was obvious enough and especially if it became established early enough, it was then and generally is now an accepted grouping. If it was proposed later and in an atmosphere in which terms like the "rigorousness" of the comparative method received more and more emphasis, it became controversial or was rejected largely, again, for accidental historical reasons, this time changing fashions and the vicissitudes of the Zeitgeist.

It might be argued, of course, that this is no whim of the Zeitgeist. Rather, our methods have improved and become more rigorous. However, that this is not so is shown by two things. Classifications arrived at without these more recent methods have stood the test of time in that they are universally accepted and form the basis for ongoing work. On the other hand, it is not difficult to show that recent so-called rigorous but actually irrelevant requirements, if actually applied consistently would dismantle such universally accepted stocks as Indo-European, Uralic, and Afroasiatic.

2.4. Relationship Versus Classification

The reason for what has been called in the previous section "relativity of proof", is that the wrong question is being asked. If we ask, whether A is similar to B, since everything in the universe is similar to everything else in

some way, one never knows how similar something must be to something else (here, of course, one language or language group to another) to draw any conclusions. What is lacking is a standard of comparison, the standard being in this case other languages. But this already brings with it multilateral comparison, degrees of similarity and thus, when carried to its ultimate conclusion, a classification of languages.

Ironically, this point has often been seen, but not carried through with thoroughness or impartiality by opponents of particular hypotheses of relationship. An interesting case in point is that of Yukaghir, in relation to the Indo-European-Uralic hypothesis. Yukaghir was up to that time, and still is now lumped by many with other languages in Siberia spoken by small populations in a Paleo-Siberian group with the proviso that none of them are provably related to any other.

Paasonen in 1907, in order to refute the hypothesis that Indo-European was related to Uralic, pointed out that there were at least as many resemblances between Uralic and Yukaghir as between Uralic and Indo-European and everybody *knew* that Yukaghir was not related to Uralic. The hidden syllogism of Paasonen's reasoning might be stated as follows:

1. Yukaghir is not related to Uralic,
2. Yukaghir shows at least as many resemblances to Uralic as Indo-European does,
3. Therefore, Indo-European is not related to Uralic.

In fact, by calling attention to the numerous resemblances of Uralic to Yukaghir, Paasonen inspired Collinder (1940) and others to advance the notion that Uralic was related to Yukaghir. This, of course, does not logically exclude the idea of a relationship between Uralic and Indo-European, only it is a more distant one. For A and B to be related in this broader framework means that they are more similar in essential ways requiring genetic explanation than some other group or groups of languages. Thus, Indo-European and Uralic are more similar to each other than either is to Sino-Tibetan. At this point, we do not posit a relationship between Indo-European and Sino-Tibetan until a still larger picture emerges.

Once we proceed in this way, we end up with a taxonomy, a classification involving degrees of relationship, one in its abstract structure quite similar to that of a biological taxonomy. To make it complete, we look for all the languages belonging with a specific language at a particular degree of relationship. Such an entire group we traditionally call a linguistic stock, and it corresponds to the biologist's notion of a taxon.

An essential difference between the approach through relationship as against that through classification is that comparison in the former case might be called de-contextualized. We compare only A and B as though no other languages in the world existed. Hence, no standard of comparison exists leading to the difficulties that have just been seen.

3.0. The Orthodoxy of the Classification Approach

The superiority of the classification approach should be obvious. From classifications, we derive numerous statements of relationship of various degrees, but it is difficult to see nor has, I believe, anyone shown how we can derive a taxonomy solely from statements of relationship. Moreover, a taxon at any level is a natural subject for comparative study and reconstruction. The proto-language thus reconstructed represents a real historically valid entity. It implies a population who spoke it, and we can derive important information about human history from hypotheses regarding its place of origin and from the reconstructed forms indicative of the culture of the speakers.

Contrary to some contemporary opinions, this has always been the basis of traditional comparative linguistics. Indo-Europeanists did not take two branches at random, say Slavic and Italic, and make a restricted comparison between them to reconstruct a Slavic-Italic proto-language. On the other hand, there does exist a field of Balto-Slavic studies because of the fairly large number of features common to Baltic and Slavic compared to other Indo-European languages. This leads to the positing of an intermediate Balto-Slavic node between Proto-Indo-European on the one hand and Proto-Slavic and Proto-Baltic on the other. Buck (1933), in his comparative grammar of Greek and Latin, makes it clear from the outset that there is no linguistic reason that would justify such a comparison, only the cultural importance of Greek and Latin literature and the fact that study of the two of them is commonly found in a single department of Classics.

The approach through multilateral comparison is merely an attempt to make explicit what the assumptions were that led to the first and universally accepted classifications into linguistic stocks. It is an effort, to use the terminology of the logical positivists, to explicate the notion of linguistic stock and family tree of languages. Such an explication involves the elimination of de-contextualized isolated comparisons as described above and its replacement by systematic multilateral comparison in order to produce a valid taxonomic scheme of languages.

4.1. Regular Correspondences as Proof of Relationships

We now consider the two most frequently used criteria, as mentioned above, which are supposed to prove relationship, namely, (1) regular sound correspondences and (2) the reconstruction of proto-languages. In fact, these two criteria are ultimately related. Textbooks of comparative linguistics, in their methodological chapters, start with sound correspondences and then show how, by using them, one can reconstruct the ancestral language. Viewed in this light, they are part of a larger whole, the initial and final stages respectively of the comparative method. Hence, one often encounters the statement that the validity of some particular stock has been demonstrated by the comparative method.

Nevertheless, these two criteria deserve separate treatment. Regarding sound correspondence, the detailed treatment in the first chapter of Greenberg (1987) *Language in the Americas* can be summarized as follows. The statement that languages are related if they show regular sound correspondence is not so much denied as shown to be irrefutable in specific cases because there are so many alternative explanations of the sort universally accepted in historical linguistics, the most important probably being analogy. Here, individual linguists may differ in regard to preferences for particular types of processes as explanations. There are those who believe that sound laws have no exceptions and who are, therefore, likely to favor regular sound changes to account for the differences in forms which appear similar enough to be likely etymologies. On the other hand, there are those who refuse to multiply sound correspondences and either reject certain etymologies accepted by the first group, or resort more frequently to other explanations including sporadic assimilation and dissimilation, metathesis, and of course morphological analogy more frequently than the first group. These are, of course, tendencies rather than two clearly defined camps with fundamental methodological differences.

There is a further point which has been seen and accepted for some time now by many historical linguists, namely that where there have been extensive borrowings from one language into another, the two languages can show great regularity of sound correspondence, indeed, often greater than between cognate languages. A now famous case is that of Wolfenden, who assigned Thai to Sino-Tibetan on the basis of what is now generally accepted to be borrowings from Chinese into Thai. His thesis is now universally rejected, but it is not always realized that in a broader classificational attempt it would be unacceptable, not only because Thai basic vocabulary is hardly affected but because every Thai resemblance to Sino-Tibetan points directly to Chinese as a source. If it were really Sino-Tibetan, it would in its vocabulary frequently display resemblances to languages like Tibetan and Burmese but not found in Chinese.

4.2. Reconstructions as Proof of Relationship

The notion of reconstruction easily derives from that of sound correspondence. Suppose that among five related languages there is correspondence such that four have *p* and one has *f*. If we take the further step of assuming that the original sound was **p*, we are making an additional, and really more venturesome assertion, namely, that the original **p* became *f* in one of the five languages. Suppose there were another correspondence in which all of these same five languages had *f*. Since the reconstruction of **f* here would be overwhelmingly favored, the reconstruction of **p* rather than **f* for the first correspondence would be strongly preferred even if only one of the languages had *p*. This would be for a synchronic typological reason. In reconstructing **p* for *p-f-f-f-f* and **f* where all five languages had *f*, both synchronic and diachronic reasons would be operating. Diachronically, $p > f$ is very common and could have occurred in four of the five languages. Further, if, in going with the majority, we reconstructed **f* for *p-f-f-f-f*, we would have two *f*-phonemes in the proto-language, which is synchronically impossible.

There are a fair number of instances, e.g., Blackfoot in Algonkian, Cherokee in Iroquoian, and Kamchadal in Chukotian, in which reconstructions have not been possible for the languages indicated, yet everyone admits that they belong to these respective families. Given this and the possibility of reconstructing a perfectly good proto-system in some instances for loan words, reconstruction becomes neither a necessary nor a sufficient condition for the languages to be related.

Strictly speaking, I believe that reconstruction in cases such as Blackfoot and Cherokee is possible by positing a very large and typologically implausible number of phonemes for the proto-language.

In an interesting article by George Grace, a leading Austronesianist, in the volume on *Linguistic Change and Reconstruction* (1990) edited by Baldi, Grace discusses two Melanesian languages of New Caledonia, universally accepted as Melanesian and hence also Austronesian which are quite closely and obviously related to each other. However, a comparison of these two languages, Canala and Grand Couli, shows 140 consonant correspondences and 172 vowel correspondences, not reducible by the existence of conditioning factors. The construction of a proto-language for these two languages would result in a language of 312 proto-phonemes. This leads Grace to the following conclusion (1990:171),

> It seems that, at the least, the question of how much regularity there really is, and under what circumstances, deserves to be re-examined on the basis of evidence.

That such an *ad hoc* reconstruction "proves" that this language is Melanesian is obviously absurd. We knew it all along. The least plausible alternative is to reject this language as Melanesian.

In one well-known instance, the vowel system of Uralic, actually that of the initial syllable, where a full set occurs, there have been two major theories, that of Steinitz and that of Itkonen. The former makes the system like that of Proto-Ostyak as he reconstructs it and the latter quite like that of Finnish. Steinitz posits a whole series of *Ablaut* variations in the proto-language, which saves the regularity by assuming inherited vowel variations, while Itkonen is more tolerant of irregularities, but in either case a large number of instances are irregular. At one point, Décsy said that in Finno-Ugrian the second part of Voltaire's famous *bon mot* concerning etymology holds, namely that the vowels count for nothing.

Since, as we have seen, Finno-Ugrian was discovered even earlier than Indo-European, and no one doubts its validity, the question arises as to whether a lack of regular sound correspondences could ever disprove a relationship based on numerous and obvious similarities. The answer seems to be that it could not. In principle, there should not be any difference between consonants and vowels in this regard. If there was ever a case for rejecting a hypothesis on the basis of irregular correspondences it would be Finno-Ugrian, but it is universally accepted. A theory not disprovable by any data is not an empirical theory. One should add that after a century or more of comparative work the consonants, especially the sibilants, present problems also.

5.1. The Internal Logical Structure of the Comparative Method

In the above discussion, there has been an implicit assumption apparently shared by all historical linguists that there exists either a set of sound laws or a valid reconstruction that is, as it were, timelessly true. We need to distinguish here two things. One is the real history of investigations by historical linguists as they take place over time. The other is *logical* priority of certain steps in the process in relation to other. We have already seen something of the notion of logical priority in the discussion of the relation of sound correspondences to reconstruction. In fact, the logically prior is likely to be (occur) first historically. Sound correspondences are noted before proto-phonemes and reconstructed forms are set up. The two kinds of priority both figure in the following statement of Delbrück, the collaborator with Brugmann in the great comparative grammar of Indo-European:

> Since obvious etymologies are the materials from which sound laws are drawn, and this is material [which] can always be increased or changed, therefore new sound laws can continually be recognized and old ones transformed. (1988:47)

First, Delbrück is obviously portraying the development of linguistics over time. New sound laws arise and old ones are changed. The initial sentence, however, contains the key notion that sound laws are in the first place drawn from obvious etymologies. But such obvious etymologies have then a kind of logical priority and must somehow be recognized by properties not involving sound laws since sound laws are derived from them.

Others have made essentially the same point. Goddard (1975:25) says,

> In general, the establishing of phonological correspondences goes on within a family of languages known to be related ...,

quoted with approval by Campbell and Mithun (1979:52).

Somewhat earlier, Newman (1970:39) had similarly stated,

> The proof of genetic relationships does not depend on the demonstration of sound laws. Rather, the discovery of sound laws and the reconstruction of linguistic history normally emerge from careful comparison of languages known to be related.

Regarding Newman's observation, Watkins, a distinguished Indo-Europeanist, has said more recently (1990:292),

> As to the mystique of sound laws on the one hand, Greenberg is quite right to quote with approbation the Africanist Paul Newman (1970).

None of these citations face the problem of the initial step, namely how we recognize "a family of languages known to be related" (Goddard) or "languages already presumed to be related" (Newman). What I believe does emerge here is the recognition of at least three stages, which are in order of logical and usually historical priority the following:

1. Recognition of a family of related languages;
2. The discovery of some sound correspondences;
3. The application of the comparative method, starting with the sound laws, leading to the reconstruction of linguistic history and of the proto-language.

In this process, sound laws are often revised or abandoned and new ones found. A core of basic etymologies furnishes the starting point (cf. Delbrück above), and it is rare for any of these to be abandoned, but also new ones are found, and some which are proposed become matters of debate with some linguists accepting and some rejecting them. Etymology will never cease to exist as a field of study in any linguistic stock simply because no more etymological problems exist.

It is a corollary of the fact that linguists' conception of sound laws and the proto-sound-system of a particular stock changes over time that a particular reconstruction cannot prove the validity of a stock. In 1891, in a joking mood, von der Gabelentz remarked that from the time of Schleicher to the time of Brugmann, Proto-Indo-European had changed quite a bit! Since the discovery of Hittite and the other Anatolian languages about 1915, the changes have been just as drastic. With laryngeal theory we now have, in the opinion of some Indo-Europeanists, only one original vowel, whereas in Brugmann there were 11 vowels plus syllabic sonants and a large number of diphthongs. If Brugmann's reconstruction was really incorrect, how could it have proved the validity of Indo-European as a stock? In my view, Schleicher, Brugmann, and laryngeal theory should be viewed as successive approximations.

We will never discover a document in Proto-Indo-European which will give us the "correct" system. Were such a sensational document discovered, reconstructionists whose views were not confirmed, could quite justifiably argue that the newly discovered document is not Proto-Indo-European but a related language or an early extinct branch of Proto-Indo-European.

5.2. The Proto-Comparative Method

By the term "proto-comparative method", I do not mean the total method of reconstructing a proto-language, but rather an initial stage of the comparative method which logically and in actual practice but without explicit recognition, precedes what is usually taken to be the initial stage in the comparative method, namely, the discovery of sound correspondences.

The pre-sound-correspondence stage was illustrated briefly in Greenberg 1987 (especially 24-26). There, in table 7, equivalents of nine common words are given for 25 languages in Europe. By the time one has gotten to the second word, a three-fold division into Indo-European, Finno-Ugrian, and Basque and even the major subgroupings of the first two (the third being an isolated language) become apparent. In comparing, for example, the third item, equivalents for the English word "three", we are making many judgments of differential phonetic similarities, the meaning here being kept constant. For example, we judge that Welsh *tri* is more similar to Italian *tre* than either is to Finnish *kolme* and in that, in turn, Finnish *kolme* is more similar to Estonian *kolm* than either is to the Welsh and Italian forms. Even these four equivalents for "three" give us a grouping of Welsh and Italian (Indo-European) as against Finnish and Estonian (Finno-Ugrian). But how do we arrive at these judgments? In doing so, we have, as it were, applied the comparative method in embryonic form. We compare the *t* of Welsh *tri* with the Italian *t* of Italian *tre* and not with the *e* in the latter word. So also, the two *r*'s and the vowels are being compared. This set of equivalents for "three" and the groupings it gives are already important evidence leading to the correct classification, and, of course, there is no recurrence of sounds within the words so there is no sound correspondence of *t* to *t* or *r* to *r* in the usual sense. When, however, we have examined more words, such correspondences will be found, and the *t* and *r* in the words for "three" will be examples of them. We might, therefore, call them proto-correspondences. Note that even after finding recurrences, we do not yet have reconstructed proto-sounds. As we have seen, this requires further steps of reasoning and is logically later.

There is further in the same section of Greenberg (1987) a discussion of the equivalents for the PIE word for "tooth" as leading back quite naturally to a single form that might be characterized as approximately **dant* or **dent*. This once again is based on exiguous data, without yet assuming anything like a complete original sound system or a regular set of sound changes. Such relatively amorphous hypotheses are the actual and the logical precursors of more fully elaborated but never conclusive formulations resulting from the application of the comparative method. They are the true first steps in the method, never recognized in textbooks of comparative linguistics but worthy of fuller study and consideration.

A further question arises from these considerations regarding reconstruction. Since, as we have seen, reconstructions both change over time and no two comparative linguists agree completely with each other at any particular time, then some, or possibly all, might be false. Does a false reconstruction "prove" the validity of a linguistic stock? The question answers itself.

5.3. Genetically Relevant Criteria

In all the preceding discussion, there have been frequent references to similarities and degrees of similarity among languages, e.g., that English is more similar to German genetically than it is to French. However, there has been no explicit treatment of the question as to what kinds of similarities are involved. After all, languages can be like or unlike each other in innumerable ways. However, there has been an implicit answer to this question in the immediately preceding section, similarity in specific word forms, or more exactly morphemes, involving sound and meaning simultaneously. Resemblances of other kinds, e.g., in sound only as in the existence of glottalic ejectives in both languages, or meaning only as in agreement in a tense system including a future versus non-future opposition, or in word or morpheme order, are irrelevant and would normally be considered typological. Of course, a language can have typological agreement with other languages of the same stock which is part of their common inheritance from their common linguistic ancestor. However, such characteristics do no figure in our carrying out the classification. They are, as it were, a bonus deriving from it and useful in considering the relative stability of typological traits and their historical development.

However, as self-evident as it might seem that it is from similarities in sound in meaning simultaneously that one finds the materials for genetic classification, there is a strand of thinking in keeping with the strong

tendency towards formalism in linguistics to dismiss phonetic similarity. All that would count would be regularity of correspondence in abstraction from phonetic content. Of course, semantic similarity cannot be suppressed. For some, it is narrowed to semantic identity, which in practice can only mean translation equivalence. The reason for this is that they wish to make it as difficult as possible to prove that anything is related to anything else.

I earlier thought that the importance of phonetic similarity was so obvious that it required no defense. We have seen in the previous section how differential similarities as in Welsh *tri*, Italian *tre* versus Finnish *kolme* "three" play a key conceptual role at the very initiation of comparison leading to classification. Those who oppose this obvious and sensible approach emphasize the frequency of accidental resemblances and the existence of phonetically unusual correspondences. Let us consider examples of each in turn.

Meillet (1958:92) gives the following frequently cited example. French *feu*, in spite of its phonetic resemblance to German *Feuer* "fire" is not cognate with it, while German *Feuer* goes back to an original represented by words which are quite different, Greek *pûr* and Armenian *hur*. The historically correct result arises from the use of the comparative method. The incorrect comparison of French *feu* and German *Feuer* comes from using mere surface resemblance. What Meillet fails to consider is how we know in the first place that French *feu* is only accidentally similar to German *Feuer*. In other words, the term "comparative method" is, as it were, hypostatized and not resolved into the concrete steps involved in its application.

We know that *feu* is not cognate with German *Feuer* because if we compare French *feu* with the Spanish form *fuego* and Italian *fuoco*, we see that it goes back to a form in which the second consonant is a velar. Compare also French *peu* "few" with Italian *poco*. Why do we compare French with Spanish and Italian in this and other instances? A glance at table 7 in Greenberg (1987:24) will show why. From the very beginning, in word after word French aligns with Spanish , Italian, and other Romance languages, while English lines up with German, Dutch, etc. Besides this, in Meillet's example, German *Feuer*, English *fire*, etc. are not all that different from Greek *pûr* and Armenian *hur* since $p > f > h$ is a very commonly encountered form of change, and they agree in their final *r*. In other words, the generally accepted cognates also show a high degree of phonetic similarity.

A favorite in recent discussion has been the correspondences between Greek *dw* and Armenian *rk* as seen in the word for "two" and a few other words. All this shows is that phonetically improbable resemblances may occur in correspondences. However, they are rare. No similar instance to *dw* ~ *erk* has ever been found, but *t* ~ *d* or *t* ~ *t* are commonplace.

More importantly, it has nothing to do with classifying Armenian as Indo-European. Scores of phonetically obvious resemblances between Armenian and other Indo-European languages exist and were noted earlier. Indeed, Hübschmann, the founder of modern Armenian studies, who first established the affiliation of Armenian on a firm basis, does not have this correspondence in his seminal paper on the position of Armenian (1875), what Meillet was later to call the "bel article" on this subject, and he (i.e., Hübschmann) never accepted this correspondence.

6.0. Biological Analogy

Beginning in the nineteenth century with Lyell and Darwin on the biological side and Schleicher on the linguistic, the basically similar structure of biological and linguistic evolution has been noted. In both instances, differential similarities result in hierarchical groups which are explained by historical development from no longer existing ancestral forms, whose distance from the present mirrors the degrees of difference. Primates had a common ancestor later than mammals, of which they are a branch, just as Germanic languages have an ancestor, Proto-Germanic, which is later than Proto-Indo-European, of which it is a changed form.

Yet, in biology, as I noted (1987:34), nothing is equatable with sound correspondence. The same observation was made by Dyen in the same year (1987:708), when he states that in biology there is no clear analogy to the law of regular phonetic change. However, he seems to view this as just one point, along with numerous other similarities and a few differences between linguistic and biological evolution without drawing any conclusions from it.

I believe that in the context of the present paper, since biological classification is not only possible but is probably, if anything, more advanced than linguistic classification, what we see is that in the very initial steps of the proto-comparative method, before regular correspondences enter in, we already have a valid genealogical classification.

The existence of regular sound changes enables linguists to reconstruct, but reconstruction is not necessary in order to ensure the validity of the classification. Once more, biologists do not reconstruct, and it is not a part of their taxonomic theory, nor does this imply any inferiority in biological taxonomic endeavors. I am confident that an application of the methods outlined here and elsewhere will ultimately lead to a taxonomy of languages not inferior to that of the biologists in their field.

References

Bopp, Franz. 1816. *Über das Conjugationssystem der Sanskritsprache in Vergleichung mit jenem der griechieschen, lateinischen, persischen und germanischen Sprache.* Frankfurt: Andreäsische Buchhandlung.

Buck, Charles Darling. 1933. *Comparative Grammar of Greek and Latin.* Chicago: University of Chicago Press.

Callaghan, Catherine A. 1958. "California Penutian: History and Bibliography." *International Journal of American Linguistics* (*IJAL*) 24:189-94.

————. 1991. "Review of Wintun Dictionary by Harvey Pitkin", *University of California Publications in Linguistics*, 95, in *IJAL* 57:129-31.

Campbell, Lyle and Marianne Mithun. 1979. "Introduction." *The Languages of Native America*, eds., Lyle Campbell and Marianne Mithun, pp. 3-69. Austin: University of Texas Press.

Collinder, Björn. 1940. *Jukagirisch und Uralisch*, Uppsala Universitets Årsskrift, 8.

Donner, Kai. 1901. "Die uralaltäischen Sprachen". *Finnisch-ugrische Forschungen* 1:128-49.

Dyen, Isidore. 1987. "Genetic Classification in linguistics and biology." In George Cardona and Norman Zide, eds. *Festschrift for Henry Hoenisgswald*, pp. 101-08. Tübingen: Gunther Narr.

Goddard, Ives. 1975. "Algonkian, Wiyot and Yurok: Proving a Distant Relationship". In M. Dale Kincade, Kenneth L. Hale, and Oswald Werner, eds., *Linguistics and Anthropology*, pp. 249-62. Lisse: Peter de Ridder.

Greenberg, Joseph H. 1987. *Language in the Americas.* Stanford University Press.

Hübschmann, Heinrich. 1875. "Über die Stellung des armenischen im Kreise der indogermanischen Sprachen." *Zeitschrift für vergleichende Sprachforschung* 23:5-42.

McRobbie, Z. 1986. "Review of English translation of Sámuel Gyarmathi's *Affinitas* by Hanzell." *Ural-Altaische Jahrbücher* 58:159-60.

Meillet, Antoine. 1958. *Linguistique historique et linguistique générale.* Paris: Champion.

Newman, Paul. 1970. "Historical sound laws in Hausa and Dera (Kanakura)", *Journal of West African Languages* 7:39-53.

Paasonen, Heikki. 1907. "Zur Frage von der urverwandschaft der finnisch-ugrisch und indoeuropäischen Sprachen." *Finnisch-Ugrische Forschungen* 23:13-30.

Pitkin, Harvey and William Shipley. 1956. "A comparative study of California Penutian. *IJAL* 24:174-88.

Ravila, Paavo. 1935. "Die Stellung des lappischen innerhalb der finnisch-ugrischen Sprachfamilie." *Finnisch-ugrische Forschungen* 23:20-65.

Watkins, Calvert. 1990. "Etymologies, equations, and comparanda. Types and values, criteria for judgment. in Philip Baldi, ed., *Methodology in Linguistic Change and Reconstruction*, pp. 289-304. Berlin: Mouton de Gruyter.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

BOOK REVIEWS

The Emergence of Homo Sapiens and His Languages in Tropical Asia, by W. Wilfried Schuhmacher, Juan R. Francisco and, F. Seto. Bochum: Brockmeyer, 1994. 93 pp.

Reviewed by Robert Blust, *University of Hawaii.*

Even in the iconoclastic and methodologically *laissez faire* world of long-range comparison, this little book, subtitled "Linguistic evidence for a Hoabinhian alternative to Engineer Anderson's thesis", must be regarded as a teratism. It contains just five pages of text including its one-page Introduction. The remaining 88 pages consist of linguistic comparisons ranging over at least 11 language families on four continents, topped off by five and one half pages of references which (1) are simply listed without any evident principle of organization (e.g., alphabetization, date of publication, order of mention), and (2) include only sources of language material — in some cases languages from which no more than one or two words are drawn for comparison — without a single reference to support any of the startling and often quite abrupt and disconnected statements which crowd the few pages devoted to any kind of exposition.

What is the thesis and what kind of evidence is offered for it? Since the linguistic comparisons do not speak for themselves, the five pages of text are critical to understanding the point of the book. But here we are left gaping into a void, as these few pages serve up what might not inappropriately be described as "word salad". I cannot possibly do better than to cite a representative passage or two. The Introduction begins (p. 2)

> Global lexical interlinkages necessarily require dispersal from a territorially limited environment. Otherwise, either the lexical interlinkages could **not** exist or there would be exceptions to them. Such a territorially limited environment prior to the dispersal of >>homo sapiens<< could possibly be only of two kinds:
>
> 1. Insular (= the **volcanic** Sumatran highlands). African comparative botany has shown that **speciation** depends on the availability of trace elements, and this is exactly what Africa is poor in. 2. Of such a naturally superior quality of life for a gathering life-style in relation to the neighboring regions that >>homo sapiens<< would have deemed it absurd or even suicidal to abandon it. There is nothing of the sort from West Bengal to the Mediterranean. The particularly wretched stony desert of Iran, with bone-chilling wintertime gales, could never have been freely opted for.

We are then told, in rapid succession and without logical transitions of any kind: (1) that African faces and ears are "fundamentally different" from those of Papuans, Indians, Middle Easterners, and Europeans, (2) that sickle-cell anaemia does not occur in non-Africans, and that the latter have, therefore, never been in Africa, (3) that Guyanese of African ancestry congregate in the cities and leave farming to the descendants of East Indians, (4) that the Mongoloid physical type arose independently in Eurasia and the Andes as a result of the "differential iodine content of soil in rain-shadow areas", and (5) that the "Hoabinhian civilization" of Vietnam (they were hunter-gatherers) is linked to the West Papuan peoples of the islands of Alor and Pantar in the Lesser Sunda Islands of Indonesia if only "this culture was capable of crossing the current of huge whirlpools between Bali and Lombok in order to reach Wetar Island and possibly Timor in the east, because all the islands are visible from one another as far as Wetar."

It is hard to know where to begin with something so incoherent. To the extent that I can make out any kind of rational claim, it seems to take the following form: (1) Johan Gunnar Andersson, a Swedish mining engineer who made the first discoveries of Neolithic cultures in China, initially suggested that Chinese civilization owed its impetus and agricultural base to influences which reached it from the West (he rejected this interpretation nine years later in favor of a southern origin; for an excellent discussion of Andersson's work, cf. Chang 1977:4-5, or Needham 1984:39ff). The misspelling of Andersson's name (he is referred to throughout without explanation of any kind as "Engineer Anderson") is symptomatic of a much deeper affliction, for through the most bizarre cogitative

machinations, his claim that Chinese civilization owed its initial impetus to influences from the West is transmuted by the authors of this book into a claim or at least implication (p. 5) that "the ultimate source of >>homo sapiens<< would have been located on the climatically sheltered western side of the Red Sea prior to the 4th millennium B.C.", (2) as a "counter-claim" (to a position which Andersson obviously never took), it is asserted that Homo sapiens arose in tropical Asia, and began to spread from there about 3,800 BC, a date remarkably close to that proposed by the late Bishop Ussher!

The whole of twentieth-century human palaeontology simply does not exist for the authors of this book (what ever happened to the Cro-Magnon invasion of Neanderthal Europe?, who left the 91,000 year old modern skulls at Qafzeh cave in Israel, or who were the H. sapiens who gathered shellfish and hunted marine mammals at the Klassies River mouth cave in southern Africa between 80,000 and 115,000 years ago?). But never mind — they are linguists, and that is where their strengths must lie.

But when we turn to the linguistic material, the chaos continues unabated, First, Nakho-Daghestani (better known as North Caucasian) is compared with Ket and the extinct Kott of the Yeniseyan language family (pp. 6-16). I am not in a position to judge the accuracy or significance of these comparisons (nor, I suspect, are the authors of the book). Next, North Caucasian is compared with Uralic (pp. 17-18) and with Austronesian (pp. 19-33). We are then treated to comparisons between Cushitic and various Papuan families and/or Austronesian (pp. 39-47), between Papuan, Austronesian, Niger-Kordofanian (Yoruba, erroneously called a "Bantu" language), Dravidian, Indo-European (Divehi = Maldivian), Australian, and Nilo-Saharan. Where I have professional expertise (Austronesian) it is clear (1) that no use is made of real comparative work, including the identification of sound correspondences and the use of reconstructions, of which more than 5,000 are available in the published literature, and (2) that the authors are fundamentally ignorant of the history of the forms they compare (e.g. , p. 32, no.132, where the Indonesian word *kota* 'fortified place, stronghold' is compared with a North Caucasian form with no indication that it is a Sanskrit loan). Where forms from more than one Austronesian language appear in the same comparison, they often have no relationship to one another (e.g., Hawaiian *puna* 'spring', Ansus *panana* 'mud', Biak *pènònèk* 'dew', bān 'to wash', Sinaugoro *vonu* 'full', Kwaio *bona* 'valley, stream bed', etc.). In short, I have the usual complaint of those unimaginative enough to care about method: the authors have simply scanned dictionaries and assembled comparisons based on the method of "aha!" rather than the unfortunately more demanding and time-consuming Comparative Method. Finally, how any of this collection of random lexical observations relates to the thesis of the book (whatever that is, and I still am not certain) is completely obscure.

The words "pseudo scholarship" have been used in recent years to describe some work in long range comparison. I do not believe that research on distant genetic relationship is in itself any less or more worthy than comparative linguistic research which permits detailed reconstruction. However, it should be more than evident to any serious scholar that there is a place for pseudo scholarship in long range comparison, and probably a considerably more accommodating place than in most other branches of linguistics. Whether you like it or not, the Linguistic Society of Paris clearly had its reasons for its infamous, but all too understandable ban.

References

Chang, Kwang-chih. 1977. *The Archaeology of Ancient China.* 3rd ed., New Haven: Yale University Press.

Needham, Joseph. 1984. *Science and Civilization in China.* Vol.6: Biology and biological technology, Part II: Agriculture, by Francesca Bray. Cambridge University Press.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

Indo-European and the Nostratic Hypothesis, by Allan R. Bomhard, Charleston, South Carolina, 1995

Reviewed by Igor Diakonoff, St. Petersburg, Russia

[Editor's Note: Unlike Bomhard's co-authored book entitled *The Nostratic Macrofamily* (Mouton de Gruyter, 1994), this book is much shorter and is aimed at a wider audience. This book is much more, however, than a mere summary of the longer book — it contains a detailed history of the search for possible relatives of Indo-European, a summary of how viewing Indo-European as a Nostratic language can shed light on several problem areas within Indo-European, a long chapter on archeology and homelands, and a chapter on the problems of Altaic, Etruscan, and Sumerian. Finally, in addition to listing all of the Nostratic etymologies found in the longer book, 50 new etymologies are proposed, and new material is added from Eskimo-Aleut. The manuscript for the book under review was submitted to Mouton de Gruyter at the beginning of the year. No decision has been made by them yet about whether they will or will not accept the book for publication.]

The computer-printed copy of the book which lies before me, is probably the most important contribution to Nostratic linguistics to date. It includes a vocabulary of circa 650 Common Nostratic stems, i.e., stems present in all or some of the six linguistic families of the Nostratic macrofamily plus Sumerian; a new system of phonological reconstructions is suggested, and important new conclusions are presented. This is certainly an excellent work and is going to be a *vademecum* for all scholars interested in long-range comparisons.

A. R. Bomhard differs somewhat from Illič-Svityč and his colleagues in the phonetic reconstruction of the forms discussed. Thus, he is of the opinion that Illič-Svityč's equation of glottalized stops in Proto-Kartvelian (PK) and Proto-Afrasian (PAA) with plain voiceless stops of Proto-Indo-European (PIE) is wrong: it would involve the assumption that the glottalized stops are the least marked members of the Proto-Nostratic (PN) stop system, which is improbable. Bomhard adheres to the idea of Martinet, Gamkrelidze, and others who regard the traditional I-E **b*, **d*, **g* as reflexes of the glottalics **p'*, **t'*, **k'*. This solution certainly allows for a better explanation of the PIE (later IE) language situation(s). However, I am not so sure that all *non* glottalic stops should be regarded as aspirated; if they were, they would have developed into fricatives during the millennia; and even if we accept the idea of their aspiration, the addition of a brevier superscript *h* in square brackets to each stop makes the typesetting of a linguistic text too cumbersome; if the non-glottalized stop were actually *always* aspirated, why not tell us so once for all, and simply assign it to our memory. However, I am not sure they *were*; I rather like to envisage the situation in PIE as similar to what happened in Semitic, where the presence of "emphatic" — corresponding to the reconstructed glottalics of PIE and PN — do *not* presuppose that all the other stops are aspirated. As for the Semitic fricatives, **ṯ*, **ṱ*, **ḏ*, I am certain that the original AA phonemes were affricates **č*, **ǯ*, **ǯ̣* (instead of Bomhard's **tʸ*, etc.); this makes the independently reconstructed Proto-Afrasian (PAA) phonological system practically identical with that reconstructed for PK (p. 53); I also believe in the existence of PAA syllabic sonants — also like PK and like IE; and I also believe in labialized velars, cf. Bomhard's reconstruction of PN on p. 63; I believe in the voiceless lateral **ŝ* (transcribed *t͡ɬ[ʰ]* by Bomhard), and in only two original vowels, **a* and **ə* (> **i* ~ **e* ; or > **u* ~ **o*, depending on whether the contact is with **y* or with **w* or labialized velars.[1] I certainly do not believe in Orel's asymmetrical (!) "original" Afrasian six vowel system, which (like many other things in Orel's reconstructions) violate the rule of Occam's Razor: *pluralitas non est ponenda sine necessitatae*. When quoting AA glosses,, Bomhard also follows my system, not that of Orel. Adopting the latter's reconstruction of AA words would mean that AA had lost a number of phonemes certainly reconstructibele for Nostratic (like the labialized velars), and simultaneously developed a number of vowels absent in Nostratic — *ex nihilo*, as it were —, which does not seem at all convincing.

I do agree with Bomhard's idea that AA was the earliest language to branch off from Nostratic. He also thinks that the same is true in the case of PK and Proto-Elamo-Dravidian (PED)[2]; I simply would not know.

There are a number of minor points on which I disagree with Bomhard, or have additional suggestions to make; but before I turn to these points, I would like to dwell on a question of major importance, namely, that of Sumerian. I do not believe that Sumerian is Nostratic. First of all, there is the question of reconstructing Sumerian phonology itself. This is no easy matter. What we actually know are the Akkadian cuneiform transcriptions of Sumerian words. Thus, e.g., in Sumerian as we know it today, there are exactly as many vowels as in Akkadian. But

Akkadian vocalism was itself reconstructed (in the XIXth century) by mechanically transferring Arabic vowels into Akkadian texts. Since the Arabic *a*, *i*, *u* were obviously insufficient, the early Assyriologists also introduced an *e*, making the supposed Akkadian vocalism asymmetrical. I have recently pointed out[3] that an *e* requires also an *o*, and that it certainly did exist both in Akkadian and Sumerian: this can be seen from Greek (Hellenistic) transcriptions of Sumerian and Akkadian words. Another point, Bomhard quotes from the existing vocabularies, such Sumerian lexemes as *du*, *dú*, *dù*, du_4, du_5, du_6, du_7, du_8 . . . du_{18}. Now, could there actually exist ten to twenty exactly homonymic verbs in any real language? The fact is that Sumerian was apt to lose the final consonant of a word (especially a stop or *γ*). Thus, e.g., *di* may actually be < **diγ*, etc. The final consonant usually reappears before a vowel morph. Also, some of the supposed *du*-stems may actually be [do]. However, this still leaves too many homonyms: these, no doubt, were differentiated by tones[4]. This is a fact of capital importance. Of course, the number of tones is unknown to us, and hence the Sumerian transcriptions tend to be unreliable for the linguist. — Also some of the Sumerian consonants have been reconstructed by modern philologists, and their reconstructions are of different value (NB the consonant **dř* is certain). But most of the Sumerian consonants are known only from Akkadian transcriptions. For example, the grapheme *ZV*, which is identical in Sumerian and in Old Akkadian (the period when living Sumerian could be heard and recorded), could stand for Akkadian [zV] (which, again, may or may not have been pronounced [d͡zV]); but it could also stand for [t͡sʔV], [t͡sV], and even, perhaps, for [sV]. Therefore, we simply do not know whether the Sumerian 2nd person pronouns *za*, *zu* can or cannot be ranged with the *t*-pronouns of the Nostratic languages. As to the 1st person pronoun, Bomhard quotes the Emesal form *ma* because it tallies better with what the Nostratic languages presented. However, far from being archaic, according to Boisson, Emesal is an artificial "lisping" form of women's spoken Sumerian (the lisp being, apparently, obligatory for women not in every word of the language but only in words specifically tabooed for them[5]. the normal Sumerian form of the 1st person pronoun is *g̃á* (NB: *g̃á-e* is originally the ergative form; I cannot agree with Bomhard's opinion that the Sumerian pronouns belong to a Nominative-Accusative system).

It is noteworthy that the number of Sumerian lexemes which Bomhard quotes in his Nostratic vocabulary is considerably smaller than that quoted by him from any other, indubitably Nostratic language. In the 650 PN lexical items proposed by Bomhard, only about forty Sumerian glosses are presented, and some of these are certainly inexactly quoted: thus Sumerian *Delebad/t* (p. 131) means 'to flare up', and, specifically, 'the planet Venus'; the meanings 'to shine', 'to be radiant', are occasional and rare; *bu-i* (p. 121) is, in Sumerian, 'to meet, to fill', very rarely and "contrived", 'to know' (the usual word for 'know' is *zu*); *bar* (p. 122) is unusual for 'offspring', etc. It seems that the author had used the dictionary of Barbara Huebner and A.Reizammer, *INIM KIENGI, Sumerisch-deutsches Glossar*, Marktredwitz, 1985 sq. In this dictionary, there are more than 100 entries under *bar*, and only one of them is 'Same(n)', i.e., 'seed', a very uncommon sense for this sign; there is actually no 'Nachkomme'. This vocabulary has to be used with caution! Besides items of common usage, it includes such as are very rare indeed, and sometimes based on misreadings (by Akkadian lexicographers, or by present-day philologists). To continue: *bad* (NB: *bad-du* is < **bad-a*) is 'to be far' etc., not 'to part'; dug_4, a very common word in Sumerian, is 'to speak', not 'to strike, to kill' (the form quoted in *KIENGI* is a hapax), etc.

The Sumerian grammatical structure differs radically from that of the real Nostratic languages, even more so than can be inferred from the author's résumé (p. 113 sq.). Note that after the appearance of the Sumerian grammar by Ms. Thomsen (and the yet not completely published studies of I. T. Kaneva), we are as near to a final picture of the Sumerian grammar as can ever be hoped for; in any case, we know Sumerian better than we do a number of Afrasian languages. To sum up, the evidence for Sumerian being Nostratic is insufficient. NB: Sumerian as a living language died out in the 2nd millennium BC; after that, it was only spoken (especially in Nippur) by a small group of educated persons (cf. p. 35). — There is no similarity between the verb morphology of PIE, or Proto-Uralic (PU, cf. p. 73), or any other Nostratic verb morphology, and that of Sumerian.

Speaking of Sumerian, which is an Ergative-Absolutive language, I would like to point out, that in Russian linguistic scholarship, it has been usual to regard the Ergative-Absolutive grammatical structure as archaic, and the Nominative-Accusative as innovative. Hence, it is interesting to have a look at the situation in the Nostratic families. Of these, PK, Chukchi-Kamchatkan, and Eskimo-Aleut are Ergative-Absolutive, but PAA, PED, PU, Proto-Altaic (PAlt), and PIE are Nominative-Accusative. It looks as if the first group were more archaic, and had lost contact with the core Nostratic languages at an earlier date. Are Chukchi and Eskimo actually Nostratic? PK certainly is, but it may have adopted the Ergative construction under the influence of the other Caucasian languages which belong to

a different macrofamily. PK is certainly strongly influenced by a Caucasian substratum and adstratum.

A few Nostratic etymologies suggested by Bomhard (no more than 5%-10%) are unacceptable to me, but space does not allow me to discuss these. I will limit myself to a couple of examples. Thus, in nos. 195, 632, and 634 (pp. 205-206), the author seems to regard the phonemes *$\underline{h}$ and *$\underline{h}$ (*x) as interchangeable. However, *$\underline{h}$ and *x have merged only in some late Semitic languages (Hebrew, Aramaic, Tigre); certainly not in Egyptian; and one cannot see why a Tamil form with $\underline{l}$ is introduced in this context. The author refers to the theory of Colarusso (p. 74), according to which a triconsonantal Semitic root is to be divided not into an original C^1C^2-element, and a secondary C^3, but into a secondary C^1 and an original radical C^2C^3. To any linguist versed in Semitic, this is obvious nonsense, and should not be quoted. There do exist Semitic verbal roots where C^1 is secondary, but they are few in number, and only a limited number of consonants (originally prefixes) can occur as C^1. — In discussing Semitic languages, the author does not mention Eblaite, which is at least as archaic as Akkadian or more so.

Being by origin a historian, I am naturally interested in the problem as to what kind of historical conclusions can be drawn from the Nostratic linguistic reconstructions and as to what kind of society spoke Nostratic (this could also help to date this proto-language). For this, I must select Proto-Nostratic terms for tools, domesticated plants, domestic animals (if any), housing, kinship terms, etc. — To my great surprise and chagrin, I found that about 90% of the glosses registered by Bomhard as Proto-Nostratic were verbs; a very small number of pronouns, and less than 10%, nouns. From my university teachers, I had received the belief that nouns were primary in language, and verbs secondary. One lives and learns. However, among the nouns quoted by Bomhard, only very few had any bearing on the problem which interested me. But anyway, here are my results:

No. 24 (p. 24): **bar*, **bər* 'seed, grain'. Not reliably attested.
No. 27 (p. 124), No. 543 (p. 180): **bay*, **bəy*, 'bee, honey' (PIE, PAA Eg); **madw* 'honey' (PIE, PU, PED)
No. 43 (p. 127): **pVʕ(w)*, **pVʕw$\underline{r}$* 'fire' (PIE, PK, PAA [Eg]; but the PED, PU, and PAlt forms, although they may be related, are less certain).
No. 55 (p. 128): **p$\underline{l}$(+C)* 'settlement' (PIE, PU, PED, PAlt). NB: not PAA!
No. 253 (p. 148): **kab*, **kəb* (PAA) 'hoof' **kap-r-* (PIE) 'he-goat, buck' (*-r-* is a classificatory morph, designating domestic animals).
No. 370 (p. 162): **ʕaw-* '(herd of) sheep (or goats)' (PIE, PAA [Eg], PU)
No. 430 (p. 168): **ʔat(t)-* 'father' (PIE, PED, PAA [EG *it*])
No. 439 (p. 169): **ʔam(m)-*, **ʔəm(m)-* 'mother' (PIE, PAA, PED, PU, PEskimo [Yupik]).
No. 454 (p. 171): **ʔan($\underline{i}$)-* 'elder female relation' (PIE, PAA [South Cush.], PAlt., PEsk).
Add. (p. 192): **dək/g-* 'earth wall' and similar; also verbal forms (PIE, PK, PAA 'to pound soil' and similar).
No. 647 (p. 212): **q^{w}ad* 'wooden vessel; gourd, calabash' (PAA, PED [Tamil]).

One would be inclined to add some more of pertinent glosses listed by Bomhard, but here one encounters difficulties. E.g., No.319 (p. 156): **k^{w}alp'-* 'dog'; this is all right for PIE **k^{w}elp-* 'whelp, puppy', but not so good for the supposed PAA **k^{w}alp'-* 'dog': the most typical representative of this root, Semitic *kal-b* 'dog' has no labialization; moreover, the *-b* (< **p'*) is a morph denoting wild animals: the Semitic term was originally applied to the wild dog!

The glosses which we have quoted here point to a settled society which had not yet domesticated the dog but was acquainted with fire and with sheep-rearing (but not cattle-rearing); not so certainly with regular agriculture; it was hardly (or but just) acquainted with pottery. This means Mesolithic, or Late Paleolithic, which is as could have been expected.

Addenda

Ad p. 29: Phrygian is not the ancestor of Armenian, but is (distantly) related to it.

Ad p. 63: it is suggested that Afrasian high allophones merged into **ə*; I think, on the contrary, that **i* and **u* developed from **a*, depending on the surrounding phonetic environment.

Ad p. 86: "Proto-Circassian" etc. must now be checked against the monumental vocabulary of all Caucasian languages published by Sergei Starostin and Sergei Nikolaev.

Ad p. 87: G.B.Djahukyan's *Interrelation of IE, HU, and Caucasian Languages* (1967) is now obsolete.

Ad pp. 111-113: Bomhard argues that Etruscan is a very archaic offshoot from Proto-Indo-European. I think this is very possible.

Endnotes

1. Not the other way round, cf. p.63!

2. When MacAlpin first suggested that Elamite was akin to Dravidian, I was the only one to support him; his suggestion was more than coolly received by Dravidologists. I am happy that Proto-Elamo-Dravidian has received the approval of Nostraticists.

3. Sumerian phonetics and phonology were much more complicated than what may have been thought reading Akkadicizing transcriptions of Sumerian texts; see in detail I. M. Diakonoff, "Proto-Afrasian and Old Akkadian. A Study in Historical Phonetics", *Journal of Afro-Asiatic Languages* (JAAL), vol. 4,1, Princeton, New Jersey, 1994, pp. 125 sqq. — Bomhard might find some of my observations profitable. Specifically on the vowels in Sumerian, cf. p. 129; bibliography p. 125, n. 225. Also S. J. Lieberman's observations are pertinent: cf. S. J. Lieberman, "The Sumerian Loanwords in Old Babylonian Akkadian" I, *HSS* 82, 1977; id., "The phoneme /c/ in Sumerian", in: M. A. Powell and R. H. Sack, etc., *Studies in Honor of Tom B. Jones*, Neukirchen-Vluyn, 1979, pp.21-28.

4. *JAAL*, vol. 4,1, p. 123 sqq.

5. The 'male' phoneme *g̃* was tabooed for women in all words.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

The Great Human Diasporas: The History of Diversity and Evolution, by Luigi Luca Cavalli-Sforza and Francesco Cavalli-Sforza. Translated from Italian by Sarah Thorne. Helia Books. Addison-Wesley Publishing Company. Reading Mass. 1995. xiii + 300 pp., illus., maps, bibliog., index.

Reviewed by Daniel McCall, Boston, Massachusetts

This slim volume, which contains the gist of *History and Geography of Human Genes* (Princeton University Press, 1994), of which L. L. Cavalli-Sforza was the senior author, is not as intimidating as its thousand page predecessor. It is the result of interviews of the geneticist by his son, who is not a geneticist, and is therefore directed toward a wider readership, not necessarily scientists, but it has attractions for any scientist or historian who may have little technical knowledge of genetics. Specifically, it can be useful to "long-range" historical linguists and other prehistorians who may have balked at the formidable 1994 publication.

What we are offered here is not merely a popularization of an enormous compendium: we get a history of the geneticist's career, and are treated to a succinct crash course in genetics for the layman (Chapter four, pp. 74-105). Finally, there is another aspect: an activist's defense of scientific truth in a political arena that does not respect truth — a postscript of 16 pages refutes the misuse of genetics in *The Bell Curve* of R. Herrnstein and C. Murray (1994).

This review is restricted to the interweaving of interpretations of genetic data with those of linguistics and other branches of anthropology. If there are problems in the genetic discussion, that is left to geneticists to critique.

The "molecular clock" method allows calculations to be made on the basis of differences in the frequencies of a large number of particular genes in separate populations; these will indicate the length of time since separation of the ancestors of the studied populations. This "clock," like the carbon fourteen and other "clocks" discovered in the last half century, is not as precise as we would like; these special "clocks" give a certain range within which an event — in this case, separation of populations ("gene pools") — occurred.

A "family tree" for modern humans has been constructed (p. 118-95). "The biggest difference in the tree is between Africans and non-Africans" (p. 119); this represents the origin of the human species on the African continent. There were some tens of thousands of years before *Homo sapiens* ventured beyond Africa, giving considerable time for genetic complexities, due to continuous rate of mutations, to accumulate in African populations. Not all of these genes were carried to other continents by the migration(s) of the earliest diaspora; there would have been a chance selection known as "founders effect." And afterwards each separate population continued producing its own mutations.

This first conclusion is significant for a discussion that had been going on for many years; it bolsters the majority position among paleoanthropologists which holds that the human species originated in Africa, but there has been a small number of dissidents who believe otherwise.

This first diaspora out of Africa, now substantiated by measurement of genetic distances, was already indicated by the presence there of our closest primate relative, the chimpanzee, with whom we share a common ancestor about 5 to 7 mya (million years ago), and by the fossil evidence of species (in our line of descent) in between: the Australopithecines, from *Australopithicus ramidus*, 4.4 mya to *A. afarensis*, 3.2 mya, which lead to *Homo habilis*, the first human (but not "anatomically modern human"), circa 2 mya, all of these extinct species were confined to Africa (p. 37-95). *Homo erectus*, who appeared in Africa shortly after *Homo habilis* (p. 43), did make its way to parts of Eurasia, but like *H. habilis*, died out, and was replaced by *H. sapiens*, whose fossil bones are found earliest in Africa — though two specimens of essentially comparable age have been discovered just across the isthmus of Suez. *H. sapiens* is two to three hundred thousand years old (p. 27), and the first diaspora out of Africa occurred about one hundred thousand years ago.

There is an uncertainty of where to place Neanderthal Man, whose bones have been found widely in Europe and more sparsely in the Middle East. Some scholars would place it close to *H. erectus*; some would make it a separate species; others, probably now a majority of the specialists, classify it *Homo sapiens neanderthalensis*. The problem is that the border between species is where infertility begins; horses and donkeys are right on the border, as we can see by crossing them to get sterile mules, but we cannot test fossils for fertility. In the early days of paleontology, every discoverer of a fossil bone wanted to give it a specific name, but as E. Mayr pointed out, there could not be that many species. Cavalli-Sforza makes the choice of including Neanderthals in *Homo sapiens*.

There is an African counterpart, *Homo sapiens rhodesiensis*, which Cavalli-Sforza didn't bother to mention here. These archaic forms of *H. sapiens* were not "anatomically modern humans," so an additional *sapiens* has been supplied, not redundantly, to distinguish us, now *Homo sapiens sapiens* from the archaic forms that still possessed some ape-like traits.

After this first split within *Homo sapiens sapiens*, caused by some of them following *Homo erectus* and *Homo sapiens neanderthalensis* out of Africa, where they eventually pushed those predecessors into extinction, there were additional diasporas from their new non-African base. Cavalli-Sforza and his associates found that:

> Non-Africans sit on two major branches. One carries today's inhabitants of Southeast Asia and the populations who most likely reached Australia, New Guinea, and the Pacific Islands from there. The groups on the other branch populated northern Asia; most headed eastward (into Siberia and then America), and the rest (mainly Europeans and non-European Caucasoids) headed westward. (p. 119)

A map on p. 122 shows probable expansion routes, and on p. 123, a table displays the data on which the map is based:

Separation of Peoples	Date	Genetic Distance
Africa & rest of world	100,000 years ago	100
SE Asia & Australia	55-60,000 years ago	62
Asia and Europe	35-40,000 years ago	48
N.E. Asia & America	15-35,000 years ago	30

The times in the column under "Date" are provided by archeology. It can be seen that the genetic distances in the third column correspond roughly to the archeological dates. Cavalli-Sforza explains:

> Genetic distance among peoples must increase with separation time, at its simplest, it will increase at a constant rate. The table shows a precise progression: as we expected, the smaller the period of time since separation, the smaller the genetic distance. Unfortunately, the [archeological] dates are approximate and, although we took averages from 110 genes, the level of statistical error remains high (around 20 percent). Taking account of statistical error, the first three comparisons agree pretty well, as if genetic distance really does increase regularly and proportionately to separation times. The last comparison is too imprecise to be reliable, although we can try to calculate America's date of occupation, starting from genetic distance and using the other three comparisons as our basis. The date obtained in this way is thirty thousand years, which falls well within those suggested by archeologists ... and is closer to the earliest limit. (p. 123)

What Cavalli-Sforza, in his polite wording, is suggesting is that the genetic calculation supports those scholars who have argued for an early date of settlement of the Americas. There is a conservative position maintaining a more recent date, though not quite so recent as the mere ten thousand years defended by A. Hřdlicka who, at the Smithsonian Institution, long presided over the debate. Juxtaposing data from different historical disciplines, as in this case, is helpful in resolving questions that a single discipline with only its own presently available data finds resistant to solution.

There are also comparisons between genetic populations and language families, both of which take time to grow, and both are due to continuous slow change after separation from one another. Basque is a language which has puzzled linguists since the early days of historical linguistics, since it resisted efforts to place it in a defensible classification with other languages in a language family. Geneticists discovered that Basque speakers stood out from

their neighbors genetically as well as linguistically. Cavalli-Sforza brings together the peculiarities of Basque in the two disciplines in an explanation which makes sense in both.

The Basque language descends from the speech of the earliest *Homo sapiens sapiens* inhabitants of western Europe. To these people, the Paleolithic cave paintings of southern France and northern Spain can be attributed. The other languages of Europe came in later, after food production (agriculture and herding) was invented about ten thousand years ago. There was perhaps as much as thirty thousand years for the Paleolithic ancestors of the Basque to develop genetic and linguistic idiosyncrasies before newer immigrants from western Asia, with their new subsistence technologies, larger social groupings, more effective weapons, different gene frequencies, and either at first, in Cavalli-Sforza's view, or as some hold, somewhat later, Indo-European languages, impinged on the ancestral Basque territory.

Before the ancestors of the Basque had moved west into Europe, they had neighbors in Asia whose languages were related to Basque. The subsequent expansion into Asia of other peoples, speaking other languages, overwhelmed most of the linguistic relatives of Basque and all but obliterated their traces on the map. Yet a few scattered languages derived from that earliest migration into northern Asia remain, although their resemblance to Basque, after so many millennia, is more meagerly attested than the evidence we find in younger language phyla to demonstrate affinity in a considerably shallower time-depth.

Cavalli-Sforza specifies the similarity of his data to the classification of language families into language phyla. As with the dating of migration to America, there is a dispute about the deep relationships among languages; Cavalli-Sforza's conclusion supports one side among linguists against the other. A section of one chapter is titled "In Defense of Greenberg" (Joseph H. Greenberg is the leading proponent of the classification of language phyla which is shown to correlate with genetic data.

Again, the geneticist's conclusion impinges on a controversy in a discipline other than his own; there are some linguists who reject the "long-range" relationships between language families, arguing that the techniques of the profession are insufficient to obtain such "deep" results. Because the genetic data correlate with Greenberg's classification, Cavalli-Sforza's research supplies a substantiation for the more venturesome group of linguists.

At several points, Cavalli-Sforza pleads for more interdisciplinary co-operation in research; for example (p. 261): "a multidisciplinary attack offers the most important insurance against mistakes." All historical work, Cavalli-Sforza emphasizes (p. 260) have a "deep basic flaw, at least for those used to dealing with the natural sciences: it lacks the support provided by experimentation." Population genetics is "historical work" just as much as the searching of archives by an historian, the classifying of languages families by linguists, or the interpretation of fossil bones by anthropologists.

Thus is it interesting that measurements of differences in gene frequencies of separate populations on all continents correlate with certain positions arrived at independently by paleoanthropologists or by linguists. In both cases, the genetic calculations find a place in disputed ground; what does this mean in terms of proof in the arguments? Genetics certainly gives more credibility to the position with which it joins, but we would be premature to close discussion — there are several regions of the world for which the gene frequencies are inadequately sampled, and — as Cavalli-Sforza's work demonstrates — additional data can alter, to some degree, interpretations of the data.

Early interpretation of frequencies of blood types (p. 113) suggested that "Europeans were a sort of intermediate group between Africans and Asians." In later years, after advances in molecular genetics, researchers in population genetics had many more genes to study, and some of them less susceptible to environmental influence than blood types, which respond to some diseases, and in the consideration of considerably greater quantity of data, the interpretation of the relationships of populations had to be modified (cf. p. 119).

More genetic research has been done in Europe than in any other region of the world; even there, more data are desirable for certain internal groupings, but particularly in Africa, Asia, and Oceania, more data are needed. The identification of Basque with paleolithic Cro-Magnon cave painters seems secure, but the relationship of Basque to certain linguistic groups in Asia may need additional collection of gene frequencies to become (if the proposition holds) more convincing. The lack of sufficient Asian data is unfortunate because the affinity of Basque to certain languages in the Caucasus Mountains, the Hindu Kush, and another "refuge area", as well as to Sino-Tibetan and one language family in North America is a hotly contested question among "long range" linguists.

Physical anthropologists might demur on Basque-Chinese common origin, but since the separation is on the order of 40,000 years, that would be enough time for the mutations distinguishing these populations to appear. The point is made that the total differences between all populations is "very small" (p. 123).

> All the peoples now living in the Tropics and the Arctic *must* have adapted to the local conditions in the course of evolution, large individual variations are not permissible in features that determine survival in a set of circumstances. We must also bear in mind that the genes that react to climate are those that influence *external features ... It is because they are external that their racial differences strike us so forcibly* ..." (p. 124)

He goes on at this juncture to refute racist interpretations, but the point here is that it is not improbable that peoples with a common origin 40 millennia ago could now be so different in "external features."

More genetic data on North China and Southeast Asia are needed:

> ... the inhabitants of Southeast Asia tend to fall with those of Australia and New Guinea. This positioning is not absolutely certain because slightly different approaches indicate that the Southeast Asians ought to be grouped with the Mongoloids who live further north rather than the inhabitants of Oceania. There are genetic variations among the peoples of Southeast Asia that the information gathered to date does not explain. (p. 120)

The uncertainty in S.E. Asia is probably due to a more recent population expansion overlying an earlier one; "In China, expansion from the first millet-producing agricultural areas was limited in the west and the north by desert and steppes." Chinese expanded from north to south, and the millet farmers mixed with rice farmers, with the result that there are "profound genetic differences between northern and southern Chinese, which reflect an ancient diversification caused by the development of two separate agricultural cultures in the north and south ..." (p. 162-63) In the last ten thousand years or less, genes from northern Asia were introduced into an ancient gene pool which had arrived in Southeast Asia perhaps fifty thousand years before Chinese agricultural population expansion began. Genetic diversity within S.E. Asia needs more research.

The diaspora which reached S.E. Asia was the earliest to move away from the first extra-African stopping place, presumably in S.W. Asia, and kept to a southerly direction because of the continuing Ice Age which made the northern areas colder and less inviting. Paleoclimatologists might make some helpful comments on the micro-ecology of southern Asian localities in the epoch. Subsequently, but not much later, Australia was reached: tools on that continent have been dated to as early as Cavalli-Sforza suggests the peopling of that land-mass.

Homo sapiens sapiens is the only non-marsupial mammal in Australia aside from a domesticated canine, the dingo, which presumably came with the humans to whom these dogs are still attached. It has been known that dogs were the earliest animal domesticated, but one would not have expected the event could be put that far back. Genetic frequencies of domestic animals (and plants) would provide a supplemental picture to culture history.

Cavalli-Sforza speculates on the forces which stimulated population expansion. *H. erectus* probably was enabled to expand its territories within and outside Africa by having superiority in its Acheulian tools, which were an advance on the Olduwan tools of *H. habilis*. *H. sapiens* was better able to communicate that *H. erectus* by having language (as well as a more diversified tool kit) and thereby spread more widely. *H. erectus* might be considered the "Homo alalus," imagined by Haekel, "speechless man" that the German disciple of Darwin intuited as the "missing link.") By fifteen thousand years ago, if not sooner, modern human had reached all continents, and *H. erectus* had been deprived of its territories and become extinct.

Between 100,000 and 10,000 years ago, the power of expansion was provided basically by improved tools and language, but food production then became a new and more powerful factor pushing faster population growth and new sets of diasporas.

These two great periods, 100,000 to 10,000 years ago and 10,000 years ago to the present, not surprisingly, differ in the quantity of archeological remains and other traces of culture — the more recent being fuller. Those in the later period should be easier to relate to historical knowledge provided by archeology, ethnology, and historical linguistics. However, this will not necessarily eliminate disputes.

Cavalli-Sforza, for example, has the Afroasiatic language family expanding from Southwest Asia; he sees it as one of three diasporas out of the neolithic heartland of S.W. Asia, generated by population growth due to agriculture, beginning about ten thousand years ago: one was Dravidian, which went through Iran into India; another was Indo-European, which went first westward through Anatolia into Europe (overrunning the linguistic relatives of the Basque while some of the Indo-European speakers turned east on the steppes north of the Black Sea into central Asia whence some of them turned south and overran many of the Dravidians); the third group were the Afroasiatics who moved south across the Suez peninsula into Africa and eventually their forward members impinged on Nilosaharan speakers helping to disperse them southward. But Joseph Greenberg, the linguist whose work Cavalli-Sforza admires, the discoverer of the Afroasiatic and the Nilosaharan phyla, has Afroasiatic originating in Africa, because four of his five members of the language family are in Africa and only one in S.W. Asia. There is no discussion of this discrepancy.

Furthermore, when archeologist C. Renfrew suggested a few years ago that the spread of Indo-European languages was carried by agricultural expansion, many linguists expressed their objections. Hittite, Lydian, and other Anatolian I-E languages were seen as intrusive from Europe rather than relics in the homeland of Indo-European, because the differentiation of Celtic, Italic, and Germanic westward was felt to be later, on the basis of the rate of expected diversification per millennia (though we have no way of measuring that precisely). These disagreements need to be worked out to the satisfaction of scholars in the relevant disciplines, but thus far, archeologist Renfrew, linguist Ruhlen, and geneticist Cavalli-Sforza are in agreement.

The call for interdisciplinary collaboration, which Cavalli-Sforza has made at several points in his career, is the way to iron out the differences in perceptions of the journey through time of our species. His efforts to expand the view from population genetics to a total history of humanity is impressive, and further discussion from him, and others in historical disciplines, is eagerly awaited.

However, to this culture historian, it is interesting in the opening chapter to see the geneticist pick his way nimbly through a mine-field of anthropological concepts. Chapter one is about Pygmies and how they fit into the social as well as genetic evolution of our species.

Cavalli-Sforza has been to Africa 10 times and taken blood from 1500 Pygmies in 30 different locations (p. 5). Pygmies are considered important (p. 24) because they continue to live as hunter-gatherers, the way of life, until a few thousand years ago, of all humans. This sounds like the "man, the hunter" syndrome, now somewhat suspect in anthropology: study of surviving hunter-gatherer societies used to be thought very significant by ethnologists for an understanding of the beginnings of culture history, but it was pointed out that:

1. The hunter cultures have as long a history as any others, all going back to the first human ancestors;

2. However slowly, it is the nature of culture to change;

3. Inuit (formerly called Eskimos) do not hunt or live the same as San hunter-gatherers (formerly called Bushmen);

4. It was shown that some East African hunters had reverted to that economy after having been cultivators;

5. It may be inappropriate to emphasize hunting more than gathering since among some hunters, the San, for example, 80 percent of their nourishment comes from gathering, though meat holds a high place in their value-system; "foragers" (includes both hunting and gathering) may be a more neutral term.

Different aspects of culture change at different rates; tools which have reached a high degree of efficiency may remain stable for long periods, but expressive features, music, art, folktales, are likely to be more effervescent [though mythologist Joseph Campbell believed he could detect paleolithic elements preserved in modern recitations]. In any such efforts our perceptions are murky.

Some hypothesized characteristics of social organization of paleolithic foraging societies have not stood up to wider observations. British social anthropologist Radcliffe-Brown and American cultural anthropologist Julian Steward both deduced that patrilocality was functionally imposed on hunters by the division of labor, inasmuch as

men did the hunting, sons would learn the ecological characteristics of the country while being taught to hunt by their fathers, and therefore would be handicapped if on marriage they joined their wife's band and had to hunt in strange country. This has since been shown to be untrue in both San and Pygmy societies.

Nonetheless, some deductions concerning early, middle and late paleolithic societies have to be attempted, but cautiously. Cavalli-Sforza, in addition to his measuring genetic frequencies of Pygmies, indulges in a few extrapolations of a cultural nature from modern hunters to ancient humans. He notes the average size of bands and the "tribe" to which the bands adhere; band size relates to efficiency in exploiting the food resources of an area, and the minimal size of a "tribe" to avoiding deleterious effects of inbreeding. So far, so good.

There seems to be an unstated premise that the genes acquired and maintained in adapting to the Pygmy life-style provide an ancient point of departure for the study of all populations, some of whom are now adapted genetically as well as culturally, to other, more recent, ways of living. But shortness is an advantage in tropical forests (p. 11) and this is not the environment in which *Homo sapiens* originated; it is generally conceded that our cradle-land was the East African highlands where trees are scarce, so in this respect, unless the earliest humans were *that* short, Pygmies do not maintain entirely the genetic makeup molded by natural selection in the original habitat. In the total picture, this is perhaps of little significance, and isolation in the forest which until a few millennia were impenetrable to others, Pygmies should have less genetic mixture than other African hunters. It may be unimportant that there has been some cultural influence on Pygmies: they now speak Bantu, having lost their own language (p. 9), and sometimes use a cross-bow, borrowed from Bantu, as well as the ordinary bow.

Cavalli-Sforza's text, which specifies such cultural influence on the Pygmies, and doesn't deny cultural change, avoids — in any statement essential to his main thrust — the problems anthropologists discovered only belatedly. However, Chapter One, "The Oldest Way of Life," skirts these problems at a few points. "The current habit of 'buying' their wives was probably borrowed from local cultivators." (p. 9) They pay with meat given to the parents of the bride. This is also the custom among the San; did the San also borrow from Bantu, or is the practice a bit of African paleolithic culture? Cavalli-Sforza in this instance has attributed to *borrowing* what the "man, the hunter" theorists would have attributed to conservative culture. In another instance (p. 7), he seems to use the "man, the hunter" interpretation: Pygmies are peaceful people. The San have also been called a peaceful people, but Richard Lee collected testimonies about personal conflicts among the San, arranged them on a scale from shouting matches to murder, and compared this record to that of a rural community in America to find that, on a per capita basis, the San hunters were not much more peaceful than the country people (in the period of the comparison).

These quibbles about Chapter One do not detract from the overall construction of the great human diasporas. The whole book is joy to read for anyone interested in the assemblage of various kinds of research that goes into the discovery of the many millennia of history of our human past.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

EDITORIAL I: SURVEYING THE DAMAGE

As editors, we have tried to be as fair, open, and honest as we could be, despite the fact that one of the three most salient and major long ranger hypotheses was under very heavy attack. We believe that we succeeded in being fair to the attackers. However, now that the debates are over — at least for this round —, we realize that in our drive to be fair to the attackers, we have actually been rather unfair to our esteemed colleague, a sturdy pioneer of our party, John Bengtson. This is not to say that Bengtson "lost" the debate, nor does it imply that his case was not strongly made. Rather, we fear that in the minds of readers, the unfairness of the presentation will lead them to wrong conclusions.

Reminding everyone that Bengtson had a number of built-in disadvantages (of which more below), we point out that Bengtson and his supporters ended up with less than 50 pages of text versus some nearly four times as many pages by his critics. Not counting about half of the supportive material, written by himself, Bengtson received about 18 pages of strong support from two colleagues (Blažek and Ruhlen). Since Blažek was also in the midst of his doctoral defense at the time, the support he was able to give Bengtson was necessarily curtailed, albeit heroic. Moreover, we had set no practical limits on the amount that could be written, which some of Bengtson's supporters took as a restriction on their output, while his opponents did not. Our reason was a desire to make the Basque discussion a thorough one.

Our esteemed Fellow in St. Petersburg wrote to express his lack of experience in Basque matters but gave us the names of two of his countrymen who might help. They were not contacted. Except for Shevoroshkin's contribution, not one of the Muscovites whom we contacted lifted a pen in Bengtson's defense, even though we had been counting on Muscovite support to balance the very heavy support we expected Trask to receive from the three vasconists. When the chips were down, the Muscovites were nowhere to be found. The lack of support received from members of the school which Bengtson has spent a decade bolstering, praising, and defending is simply unconscionable! And such red faces for the editors, who had set the whole thing up, expecting a vigorous defense of Dene-Caucasic from Moscow. Since we communicated four times with Starostin by e-mail, there is no doubt that he and his colleagues knew what was going on, what the issues were, and what they could do about it. Aside for a startling whiff of paranoia amongst them, it is hard to escape the conclusion that comrade Bengtson was just abandoned.

The second major disadvantage was that of the comparativist. Local experts tend to react protectively when confronted by schemes using their data, their "tribe", for general purposes. Anthropology is familiar with the classic stance of the British social anthropologist, whose life work might be the intense study of "his people", against the Boasian or Murdockian comparativist. Linguistics has become so specialized, albeit with an eye on high theory not exotica, that Isidor Dyen once said at a conference that "the Anglophiles have taken over!", and no one disagreed with him. There is usually a discouraging amount of literature to plow through in order to say that one is acquainted with the local scene, often written in languages not familiar to the researcher. But the comparativist must *per force* come to grips with the local expertise, not only because there is much useful to learn but also to avoid being criticized for neglecting good work.

But in a very real sense, long rangers have known about this problem since our inception as a collectivity, and we may now chastise ourselves for neglecting it. In MT-1 and MT-2, Dolgopolsky and I discussed the need for cooperation. No one scholar alone, probably not even Greenberg, could do it all. With our different expertises, we could help each other; and the comparativist would be helped by the local expert, not attacked. We have succeeded to some extent in doing this, but the Zeitgeist of contemporary historical linguistics is so inward-looking and defensive that our networks of cooperation cannot cover some large areas, especially the Americas. By accident, we have no vasconist long rangers. Our Burushaski long ranger never responded, nor did our Caucasologists, which means that scholars may not cooperate when they should. Perhaps scholars have such intense self-interest that seeking their cooperation is a futile exercise.

What were Bengtson's other disadvantages? First, even if some good colleagues disagree with us, we think that using the proto-Caucasic reconstructions of the Moscovites as a central focus caused Bengtson a great deal of trouble, not only in trying to relate Basque words to them but also in trying to follow their "system" of thinking about proto-forms. One of us has been doubting the Moscovite reconstructions for several years now, as readers

well know. In addition, the failure of one expert Caucasologist to comment on the article was due in part to his misgivings about the reconstructions. The other of us has been told by two Caucasologists with whom he consulted that they have many reservations about the proto-Caucasic reconstructions proposed by Starostin and Nikolaev. But Bengtson did trust and use that work.

Finally, Bengtson had two other disadvantages, to wit, (a) he is employed full time at a non-linguistic job and (b) he is bold and venturesome. The first trait is obvious and needs no comment. But the accusation of hazarding hypotheses is an important one in contemporary linguistics — it means that one is prepared to live dangerously. By the very nature of things, one who dares is one who is exposed. Bengtson would seem to be "fair game" in a "feeding frenzy" for ambitious younger linguists. Never fear, Bengtson is plenty tough — being burned occasionally speeds up the learning process!

But the pioneer aspect of things was perhaps not appreciated adequately by the pioneer himself. This is not a disadvantage but simply a mistake. Consider how difficult it has been to classify Basque over many generations of scholarly effort. By ordinary common historical sense, that shouts *remote* at us. Single languages which are distant genetically are always hard to classify because the evidence gets slim, and borrowing is so often a major factor. It has been said before, even here in *Mother Tongue*. Like Burushaski, Nihali, Kusunda, Sumerian, Ainu, and, yes, Japanese, Basque will probably need more time for its classification to be agreed upon. Consider one of the difficult languages also spoken by mountaineers — Armenian. Were it spoken in the Pyrenees, does anyone doubt that it still would have been classified as Indo-European long ago? Bengtson's proposed Macro-Caucasic will not be a young phylum like Indo-European — it is likely to be older than Afrasian or Nostratic or at least the same age.

Little was said, little is usually said, about Basque as a dialect cluster. Basque is arguably the equivalent of Latin or Italic in the time depth of its internal diversification. How far apart are those so-called Basque dialects? Again arguably, Germanic is a group of dialects just as much as Basque is, except that the accidents of history have created written standard forms as foci within a dialect continuum, much the same as Romance or Indic. The bearing of this observation is that the flow of loanwords into Basque is often undated. Could Basque of 100 BC borrow from French or Spanish dialects of, say, 800 AD? This question arises often during the debate but most pointedly in Jacobsen's penultimate remarks on the borrowings from Romance, which wipe out virtually all the core or most conservative lexical items in Bengtson's argument. Aside from the phonetic implausibility of some of the match-ups, which also require belief in a complicated system of phonological reconstruction, the relative ages of the contact languages becomes highly relevant.

By Caesar's time, Latin confronted Basque along a wide front in Gaul and Iberia. Might loanwords have flowed both ways, more than thought? Should we expect putative Latin loanwords to Basque to be verified by some Indo-European cognates? Perhaps most of all, an outsider, a non-vasconist, would expect that Spanish had borrowed a lot from Basque, noting the numerous apparent cultural similarities (e.g., the bull rituals), their intimacy as embattled northern Christians in Muslim Spain, and so forth. Since Spanish as a Romance language is partly defined by its copious borrowing from Arabic and Berber, where are the references to such borrowing into Basque?

On Natural Causes versus Inheritance

We wish to comment on the appeals to "phonesthetic", "onomatopoeic", "sound symbolic", or "nursery" causes or factors which were used to dismiss putative cognates. The reasoning was somewhat facile in places because the authors relied on some imagined agreement that linguists have reached on these matters, or so they suppose. We make two points. First, a reverse "generic fallacy" is evident in some of this. The generic fallacy says that, if you know the origins of some thing, then you know what it is. Since the beautiful butterfly began life as a squishy worm, then it must be a squishy worm. If a parvenu peasant rises to riches and fame, s/he can still be dismissed for being just a peasant. The *reverse* generic fallacy would be to deny origins in favor of current forms. If it is a beautiful butterfly, one denies that it could have been a worm. Glorious mankind could not be descended from apes. If a nursery word of 10,000 years ago lingers on in a language family, then it can be dismissed as just a nursery word, i.e., one made up today, made up "all the time" by children. Or a word for "fart" or "nose" cannot be old because it imitates nature, i.e., the sounds of those items.

Second, we know that much of sound symbol theory was long ago refuted in historical linguistics, at least

in the form of the "bow wow" theory of word or language origins. English, German, and many others share an old word for 'dog' and a not-so-old word for 'cat'. English has many words imitating what dogs "say": *yap*, *growl*, *woof*, *arf*, *bow wow*, *howl*, *bark*, etc. These are made up supposedly in an instant. Yet they have persisted all our lives! The cat calls of English: *hiss*, *purr*, *meow*, *growl*, *yowl*, etc. are the same. The key word in all this is *imitation* because the words do resemble some of the sounds cats and dogs make. This is well brought out by the times we try to closely approximate the real (canine and feline) sounds. Typically, they lose their consonants, suggesting that the English words have an element of *arbitrariness* in them, combined with good imitation. Yet they are also *conventions*, making them properly *icons*. Since they thus show some attributes of *symbols*, we may conclude that one language's icon need not be another's.

Two languages may have inherited such icons from a common ancestor, so that some resemble each other. English *fart* and Greek *perd-* come from PIE **perd-* and thus are cognate but sound imitative, but maybe only the [p] was originally imitative. Arabic *fasā* 'to fart' and Amharic *fässa* 'to fart' are highly imitative but also are inherited from proto-Semitic. They might connect with PIE **pezd-*, an alternate to **perd-*. If so, then English and Amharic 'fart' may be derived from proto-Nostratic of long ago. Let us state the true principle in these onomatopoetic matters:

Cognation between icons cannot be denied simply because of icon-hood, especially if the symbolic elements appear to be cognate.

Another kind of "natural cause" exists, i.e., explained by something other than inheritance. They are not strictly phonesthetic but rather retentions from early child learning, where [b], [m], [t], [d] are the easiest and earliest phones mastered. The best examples are from Indo-European. Given the alleged universals for parents, i.e., *baba* and *mama*, then one can plausibly argue that the most precious set of Indo-European cognates — the family ones — can be rejected because they contain them. So English and Spanish *fa-ther/mo-ther* vs *pa-dre/ma-dre* can be thrown out? Few linguists would agree to do that! Moreover, English has terms of address like *pop*, *papa*, *mama*, *mom*, *mommy*, *dad*, *daddy* which are more commonly dismissed because they are nursery words supposedly.

One of us (Fleming) has watched them arise spontaneously in four children. But with considerable help from care-givers they usually get consolidated into the small set of alternatives offered by English. Having heard all kinds of young Ethiopians say [baba] or [abba] for 'father', but hardly ever [dada], he wonders why he almost never hears young Americans say [baba]. It certainly looks like cultural transmission, doesn't it?

Our point is that many things of sound symbolic origin, or early learning, acquire a place in the language — and the kinship system of the culture — and subsequently are inherited by the users. They may acquire reinforcement, confirmation, and renovation from the original processes, especially as each new batch of children finds them the easiest way into the language, but they are *still inherited.* And rather than being without value in historical linguistics, we think the opposite is true; they may have great value — as survivors of proto-Human.

We want to dwell just a bit longer on this point because we believe it is important. Consider the characteristics of the human nose and its operations. One can think of many ways in which it might be depicted in language, but let us concentrate on its operations. Most of them involve air moving through the nasal cavity, coming and going. Associated with these operations are two sounds which also associate with farts, viz., [f] and [s] primarily but also [p], [z], and [ts]; broadly speaking, the flow of air or wee explosions of it. With the nose, however, we must add [n] which is a resonant quality of the nose, but also [m], [ny], and [ŋ]. If we follow sound symbolic theory, then words containing these phones, and meaning nose or one of its operations, should be dismissed. "It could be invented any time." Indeed, M. Lionel Bender and others used this argument to deny the validity of the world-wide /sun, sunt', suŋ/ etymology, in MT-2 or 3. By implication, Watkins' proposed **sn-* for proto-Germanic met the same objections. The word proposed had, of course, both an [s] and an [n], an ultra-nasal thing!

What then were we to do with English *nose* and German *Nase* and some others of that ilk? They had both <s> in its voiced form [z] plus [n]. On exactly the same grounds that Bender used in our first example, we would have to reject this cognation for 'nose'. Ah, but nobody wanted to do that. Why not? Well, the English and German forms were well grounded in Indo-European where the ancestral form (circa 5500 years ago) was a very similar: **nas-*. The problem was put back about six millennia because the heavy evidence of inheritance could not be circumvented. Moreover, we find that Amharic has a word *nəsr* 'nose-bleed' which may be cognate at the Nostratic

level — a long time ago.

The burden of proof, a need to show in detail why such and such is so (true), must be borne by the opponent of etymologies with potential natural causes, not simply by the proposer of those etymologies.

An ancillary observation about the debate over Basque is what seems to be an extraordinary amount of basic or conservative vocabulary attributed to *borrowing* from Indo-European languages. Are we to assign Basque to the category occupied by Albanian and Nahali — languages at the extreme of borrowing into basic vocabulary or at the extreme for non-retention of basic vocabulary? Since so much of the borrowing seems to be buttressed by (alleged) sound correspondences, might we entertain the hypothesis that Basque is in fact genetically related to Indo-European? Or is Basque a language so overwhelmed by its neighbors that it has no core of taxonomically useful vocabulary left? In brief — it cannot be classified?

Finally, a query to the vasconists whose strong, sometimes brilliant, certainly damaging, rebuttals to the Bengtson-Čirikba hypothesis have been vigorously pursued. Inside yourselves, why have you all been so intent on "destroying", "ruining" this hypothesis? Who is the maiden you defend so stoutly? What would have been the outcome if you had devoted just as much energy and knowledge to trying to improve the hypothesis? You have shown that your expertise in Basque is great. You can expose false statements about Basque words and grammar. And you can suggest, with considerable acumen too we must add, that all searches for similarities are foolish. But you may have also shown some myopic contentment in proving just these things. (One vasconist has reported strong social pressure to preserve the uniqueness of Basque. So, presumably, fending off comparativists brings social approbation?)

Yet in a very serious sense, you cannot say that you have *falsified* the hypothesis; for two reasons. First, you have only "demolished" Bengtson's and Chirikba's presentations of evidence. Assuming, for whatever reasons, that neither scholar presented the evidence properly, then there remains a body of evidence you have not yet destroyed because it has never been presented. Second, you have not tested the Macro-Caucasic theory because you have not taken your expert's knowledge of Basque with you to examine Burushaski and Caucasic. So we wish good hunting to you, as you take the Basque data and go looking yourselves for matches, look-alikes, similarities in the Caucasus, in Pakistan, and elsewhere.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

EDITORIAL II: AFRASIAN GETS BETTER, BETTER, AND DEEPER

The phylum of many names keeps adding to itself; new members, new insights, new reconstructions. Afrasian is Igor Diakonoff's term for what most of us used to call Afroasiatic, coined by Greenberg. Afrasian is just more apt and shorter than its predecessor. That's all. It has been called Semito-Hamitic, Hamito-Semitic, Hamitic, Erythraic, Egypto-Semitic, and other things; surprisingly some of those terms are still used by scholars. Some of the most extremely conservative and hyper-empirical scholars on earth can be found laboring away in the bowels of Afrasian antiquity, yet some of the most venturesome fly around on its surface.

Afrasian has been sneered at by at least one Indo-Europeanist, who reckoned that it was based on guess work, not solidly based like Indo-European. His point was to disparage Nostratic and other long range proposals; he didn't have it in for Afrasian in particular. But it is surprising how poorly informed he was. In phyletic terms, Afrasian was a whole lot bigger and much older than Indo-European. And even in those days — it was only 10 years ago — Afrasian was pretty well established.

Or we Afrasianists thought so anyway. But the point was important to the Indo-Europeanist because he had to prove that things bigger and older than Indo-European could *not* exist, i.e., could not be established properly, i.e., could not show systematic sound correspondences and reconstructions. Otherwise, how could he prove that Nostratic and such were impossible?

Recent work in Afrasian has gone far towards "destroying" his point. Not only have Diakonoff and his colleagues produced a proto-Afrasian of the "proper" kind (*Proto-Afrasian and Old Akkadian: A Study in Historical Phonetics* [1992, Princeton, NJ: Institute of Semitic Studies][= *Journal of Afroasiatic Languages* vol. 4, nos. 1 and 2, Fall 1992]), his junior colleagues Stolbova and Orel have produced a second one (*Hamito-Semitic Etymological Dictionary* [1995, Leiden: E. J. Brill]). Stolbova and Orel has not been sent to us yet for review.

Both of these books by eminent Russian scholars involve years of hard work, ultimately inspired by Igor Diakonoff and influenced by his teachings. It would cause one to anticipate good quality historical work, and rumor says that's true. And like most of the Russian work in Afrasian, it will be slanted towards the north, towards the eternal verities of Egyptian, Semitic, and Berber, but Stolbova's work on Chadic adjusts that emphasis somewhat. There does not seem to be any significant taxonomic change, except that Omotic now shows up in Russian comparative work — a little.

A third book by Christopher Ehret is at hand, entitled *Reconstructing Proto-Afroasiatic (Proto-Afrasian): Vowels, Tone, Consonants, and Vocabulary* (1995, University of California Press, Berkeley, Los Angeles, London. 557 pp.). Ehret has worked on this for years, anticipating it with a bold attempt to relate the prevailing triconsonantal root system of Semitic to the biconsonantal roots prevailing elsewhere in Afrasian. As is proper for someone whose first work was at the southernmost extreme of Afrasian in Tanzania, Ehret gives Cushitic and Omotic more attention than they usually get. Moreover, his sound correspondences are oriented to a different taxonomy than the usual, although he tries to keep one foot solidly planted in Semitic whilst searching for the remoter matches in Omotic. His taxonomy is interesting and different, as follows:

I. OMOTIC

- A. North Omotic
- B. South Omotic

II. ERYTHRAEAN

- A. Cushitic
  - 1. Beja
  - 2. Agaw
  - 3. East-South Cushitic
    - a. Eastern Cushitic
    - b. Southern Cushitic

B. North Erythraean
1. Chadic
2. Boreafrasian
a. Egyptian
b. Berber
c. Semitic

If his taxonomy is correct, then one can hardly credit some Nostraticists, whose Afrasian usually is little more than "Boreafrasian". Arabic and Hebrew are not just sacred languages in religion, you see. One would hope that the day when two or three Semitic languages are taken to represent the whole of Afrasian will soon be behind us. Both of us (Bomhard and Fleming) like Ehret's taxonomy, though one of us (Fleming) disagrees with some things, especially II.A.1 and II.A.3, while proposing that Ongota be entered in as either II.C or simply III. He (Fleming) usually finds himself in agreement with about 50% of Ehret's etymologies, upon which the reconstructions rest, but in this case of a deeper and better book, he agrees with maybe 30%. He (Fleming) will be pleased to review this book in MT, hopefully in the same issue with a review of Stolbova and Orel. Disagreement on many etymologies does not, of course, necessarily mean rejection of the prevailing sound correspondences, since these are usually multiply based.

An important criticism of Ehret's reconstructions, not his taxonomy, is his avoidance of Berber. As one of the traditional five (1963) or six (1987) sub-classes of Afrasian, Berber is rather too conspicuous to ignore. While it may be true that the basal sounds of Afrasian can be found without Berber, Ehret automatically invites criticism from Afrasianists by putting off a consideration of Berber. Granted that Berber phonology can be quite difficult and granted that external influence has critically affected it, still Berber is an important item of African and Mediterranean prehistory. If one believed in sub-strata, one could argue that the pre-Afrasian "elements" in North Africa — and they could be Basque-related — might be revealed by what happened to Afrasian in the Maghreb.

The implications of the new Afrasian taxonomy are serious for the twin problems of dating proto-Afrasian and locating it in space. Since II.B.2 has already within it three of the oldest dated fossil languages on earth, Archaic Egyptian, Akkadian, and Eblaite, and since II.B.2.c itself is as old as Indo-European, the antiquity of Proto-Indo-European appears paltry compared to proto-Afrasian.

Moreover, Ehret's taxonomy does not add strength to the hypothesis that proto-Afrasian was a Neolithic culture, Natufian, which moved from the Levant over and down into Africa. *Au contraire*, it seems more likely to have moved into the Levant from Africa.

Donald Levine once said that some scholars contemplated Ethiopia — and by implication the whole Afroasiatic realm — from the viewpoint of Babylon, "with the seat of their pants in Babylon". Most comparativists do the same thing. We wonder when the *fixation* on northern Semitic will lighten up, but doubt that it ever will. Ehret's book seems a great leap away from the Babylonian captivity.

MOTHER TONGUE

Journal of the Association for the Study of Language in Prehistory, Issue 1 (December 1995)

MORE ON THE AUSTRIC HYPOTHESIS AND AUSTRONESIAN'S INCLUSION

La Vaughn H. Hayes
Fayetteville, NC

[Editor's Note: The following letter was received in October 1995. Because of its length and the seriousness of the discussion, we decided to present it as an article, rather than report it in the Newsletter. Dr. Hayes is an important commentator on things Austric, whose work we have been largely unaware of, although two long rangers have recommended him to us this year. One of us (Fleming), as the alleged misinterpreter of things Austric in MT-25, does not wholly agree with Dr. Hayes' version of what we said or how we proceeded in MT-25. But both of us are very pleased to hear what he has to say and to welcome the revived discussion of Austric. We also sympathize with his irritation about not being recognized as a hard worker in a field he cherishes. Such has happened to us too. La Vaughn H. Hayes's address is: 2021 Biltmore Drive, Fayetteville, North Carolina 28304.]

I take strong exception to your Summer 1995 *Mother Tongue* newsletter commentary addressing the Austric hypothesis, which proposes that the Austroasiatic (AA) and Austronesian (AN) language families of South East Asia are genetically related. You may not be aware of it, but your commentary serves to support and propagate certain errors, misrepresentations, and false accreditations which have arisen with respect to the Austric hypothesis over the past three years. It is about time that they be addressed in a public forum.

Point 1: Questions of science must be dealt with on the basis of empirical facts, not on the personalities or reputations of scientists.

The central thesis of your commentary is that by accepting the Austric hypothesis, Robert Blust "has tipped the balance between cautious and very cautious scholars" in favor of the hypothesis because "he has great prestige as the finalizer of AN taxonomy and as a very competent but careful law-abiding 'professional'." One might observe that AN taxonomy is anything but finalized and that "law-abiding professionals" do not neglect relevant research publications when it suits their personal prejudices, but such comments border on the polemic, and I wish to avoid that tone in this commentary. The important point is that by accrediting so much to Blust's "prestige", you take the issue of the Austric hypothesis out of science and into the arena of public relations. I reject that notion as unhelpful, unproductive, and irrelevant (see further under Point 4). Confirmation of this hypothesis will be decided on the basis of the missing lexical evidence needed as the requisite proof of its validity and on nothing else (see Point 3).

Point 2: Any commentary on a hypothesis of linguistic genetic relatedness must be regarded as flawed if it does not take into account the major and/or most recent publications on that hypothesis. This is particularly true in the case of the Austric hypothesis because there have been so very few publications dealing with the Austric question.

Your commentary is based primarily on a single paper by Blust, "Beyond the Austronesian Homeland: The Austric Hypothesis and Its Implications for Archeology". It also refers to a paper by Lawrence A. Reid, "Morphological Evidence for Austric". Your commentary contains no mention whatsoever of my own article, "On the Track of Austric, Part I. Introduction" (hereinafter referred to as Austric I), which was published in *Mon-Khmer Studies XXI*, pp. 143-177. To enable you to better understand this and other points made here, I have enclosed a copy of Austric I for your examination. I hope that after reading this article, you will agree with me that it is one of the most important and insightful works ever to appear on the Austric hypothesis and should not be neglected by any serious student of or commentator on the Austric question.

In that regard, it is also noteworthy that neither Blust nor Reid mentions Austric I in their above-cited papers. Austric I appeared in print in December 1992, almost a full year before Blust presented his paper at a conference organized by Ward Goodenough and held at the American Philosophical Society in Philadelphia, PA,

during the first week of November 1993 and about five months prior to Reid's presentation of his paper at the CAMAC (Conference on Austronesian / Mainland-Asia Connections) held at Honolulu, HI, on May 7-10, 1993. Reid claims in a later paper that Austric I was ignored because it had not yet been published at the time of the CAMAC, which is untrue, of course. Why Blust neglected Austric I, I have no idea; I had advised him three years earlier (in 1990) via personal letter that the paper had been submitted to the publisher.

Point 3: Confirmation of a proposed linguistic genetic relationship is often a subjective matter, as pointed out in Austric I, but generally speaking, a certain amount and type of properly recorded and analyzed data supporting the validity of that hypothesis is required. Such data for the Austric hypothesis simply does not yet exist in print — any claim to the contrary is a misrepresentation.

In that regard, it would appear that Reid, Blust, and their supporters are attempting to change the rules of the Austric hypothesis validation game for their own personal reasons. In 1906, Schmidt presented phonological, morphological, and lexical evidence for his Austric proposal, and for nearly 90 years, the rules of the game have been that while the phonological and morphological evidence is sufficiently convincing to establish the Austric unity, the lexical evidence is not and for the latter reason the hypothesis must be rejected. That is why Schmidt has never received recognition as confirmer of his own hypothesis. As a consequence, the main thrust of Austric studies, such as they are, has always been to discover that missing, sufficiently convincing lexical evidence, and to date no one has yet been able to produce it in print. Reid, Blust, and supporters are no different in that respect. While they may ballyhoo the morphological evidence discovered by Reid as confirmatory proof, somewhere in the small print you will find their lamentations about that still missing lexical evidence which they also cannot find. Thus, any claim that Reid and/or Blust has confirmed the Austric hypothesis on any other basis is simply false accreditation.

To cite a specific example, the morphological evidence for the validity of the Austric hypothesis presented by Reid is absolutely identical to that presented by Wilhelm Schmidt in his 1906 article, "Die Mon-Khmer-Völker, ein Bindeglied zwischen Völkern Zentralasiens und Austronesiens", *Archiv für Anthropologie*, Band 33, N. F. Band 5, pp. 59-109. When the evidence presented by Schmidt and Reid is identical, how can one argue that Reid has made a convincing case for confirmation of the Austric hypothesis? If morphological evidence is to be the primary and/or sole criterion for confirmation of that hypothesis, then it was Schmidt who confirmed it in 1906 already, but in truth he has never been granted such recognition in the intervening 89 years.

Point 4: To make credible statements about the nature of Austric, in particular about confirmation of the Austric hypothesis, one must obviously have some depth of general and technical knowledge about *both* Austroasiatic and Austronesian. It does not hurt if one has also carried out some work purporting to evidenciate the reality of the Austric grouping.

Blust may be an accomplished historical linguist and Austronesianist — the same can be said of Reid, but at the time his above-cited paper was presented, Blust had done absolutely no work whatsoever which could be construed as contributing to confirmation of the Austric hypothesis. As far as I know, that is still true today. His lack of familiarity with Austroasiatic becomes obvious when one notes that he not only fails to reference Austric I, but cites no publication on Austroasiatic more recent than Schmidt's 1906 article. In essence, all that Blust has done with respect to confirmation of the Austric hypothesis, is write a paper accepting Reid's morphological evidence therefor and make some conjectures about the implications of the reality of Austric to archeology and other things. Even then, his inexperience with AA studies has led Blust to commit two egregious *faux pas* in his speculations about Austric (see Point 5). Under such circumstances, I find your assessment of Blust's role in Austric studies overdrawn and underwarranted.

Point 5: The locating and dating of "original homelands" has become a popular endeavor over the past two decades for those interested in macro-comparisons in South East Asian historical linguistics. A number of papers have appeared over the past few years in which various types of evidence are adduced in support of such locating and dating theories. On the whole, hard evidence supporting any of them is sparse.

There is to my knowledge absolutely no evidence of any kind, in or out of print, which would permit us to propose with any degree of accuracy where the original homeland of the Austroasiatic or Austric speaking people may have been. Hence, Blust's claim that the Austroasiatic homeland is to be situated in the three rivers area of southwestern China comes as a total surprise to the Austroasiaticists I know, even though Blust claims to come with the idea after consulting unnamed specialists in Austroasiatic. I wrote Blust on January 17, 1994, sending him a copy of Austric I, commenting on his 1993 paper referenced above, and asking for his reasons for placing the Austroasiatic homeland there. He has never replied to my letter. Thus, this "homeland" claim must be regarded as conjectural error, if not worse.

In the same letter to Blust, I also queried him on his dating of Proto-Austroasiatic to 5,500 BC. That dating is nothing less than ridiculous when one reviews the lexicostatistical work done within the AA field over the past 25 years. As shown in Austric I, that work indicates that Proto-Austroasiatic can be dated lexicostatistically to the 2000 to 2500 BC period. On what basis one can add an additional 3,000 years to the latter figure is simply beyond my ken.

After reading Austric I and the above comments, I hope that it will be adequately clear to you why I am simply appalled by the errors, misrepresentations, and false accreditations which have been and are still being propagated about the Austric hypothesis over the past three years, in the Austric commentary in your Summer 1995 newsletter and elsewhere. I have spent a great deal of time, effort and money over the past 13 years in an earnest attempt to discover the missing lexical evidence required to validate the Austric hypothesis, and I have made considerable worthwhile progress in accumulating such evidence (yet unprinted), a result quite in contrast to the factless conjectures by Blust and others. Nevertheless, I still do not yet consider this hypothesis to be adequately confirmed, and I doubt that I will consider it affirmed until after a significant amount of core vocabulary has been reconstructed for Proto-Austroasiatic and Proto-Austric. In any event, my decision will not be based on the acceptance of Austric as a viable hypothesis by Blust or any other *illuminati* of great prestige, and neither should your decision nor anyone else's.

[Editor's Note: Since we feel it is inappropriate for both of us to respond formally to Dr. Hayes's letter, all comments are restricted to a few to be made by the writer in the Summer newsletter, Hal Fleming.]

Hal Fleming's brief reply: There are very serious issues for long rangers in this Austric discussion. I want to skim over a few, so that our future may hold more lengthy discussions of them.

1. Lexicostatistics ≠ glottochronology; the first does not compute dates and has slightly looser rules. *Relative dating* by lexicostatistics is possible but it is all relational, not calendrical.
2. I cannot believe the dates for proto-AA proposed by Hayes. Are those languages even more obviously related than Indo-Eurpean ones? Remember that I was skeptical of Blust's dates as too young? How about somebody putting out some percentages of retention where we can all see them and make standard calculations? Furthermore, since no one believes, apparently, that much lexical evidence for Austric exists, how can one avoid the conclusion that (semantically identical) shared lexical retentions must be *very low* and the ancestor therefore *remote*? In other words, both proto-AA and proto-AN are quite young, they say, but their shared ancestor is very old, they seem to say.
3. We need to remind everyone that Austric is just about the only candidate for *autochthone* in Southeast Asia. While it is possible that some parts of that rich region were occupied by remnants from an Australian or Indo-Pacific original presence, and Sino-Tibetan lurks in the north, still Austric speakers remain the most likely general occupants of mainland Southeast Asia, especially the highlands. Given the likelihood of 40 or 50 millennia of Homo sapiens residence — at least —, should we not expect Austric to be one of those ancient super-phyla, like Amerind or Nostratic?
4. I think it is possible for two groups of languages to be proposed for a genetic connection on the basis of morphological links. It would take an old relationship to have little common lexicon remaining but it could also

be one in which borrowing had heavily depressed lexical connections. I am sure that some Semiticists, for example, would still relate Semitic to Cushitic or Egyptian on the basis of morphology (verb paradigms paramount) alone. And the converse could also be true. They tell me that Celtic could not be related to the rest of Indo-European for some time because it lacked the "solid grammatical evidence" so dear to some. Morphological evidence *is* evidence. Lexical evidence *is* evidence. With younger connections, we usually get both easily. More remote gets harder, unless there are many languages in the group. Then older means harder work for the linguist.

5. I feel much beholden to Dr. Hayes for his strong committment to work on the Austric hypothesis. I admire his tenacity and I wish him God speed in his search for lexical nuggets, wee treasures of prehistory. Had we the money we would finance his fruitful research. At least we can urge others to think about the financial and moral support they can give him. (He has not asked for anything, of course.) Send him data and ideas!

W
DE
G

Walter de Gruyter
Berlin • New York

Allan R. Bomhard and John C. Kerns
The Nostratic Macrofamily
A Study in Distant Linguistic Relationship

Trends in Linguistics. Studies and Monographs 74
1994. XII, 932 pages. Cloth. DM 298.00 (approx. US$ 195.00)
ISBN 3-11-013900-6

Mouton de Gruyter

This monograph deals with the comparison of Proto-Indo-European with various languages and language families of northern Eurasia, Central Asia, the Iranian plateau, the Indian subcontinent, the Near East, and northern Africa to determine whether or not there is evidence for a genetic relationship. The authors conclude that the evidence points strongly to a genetic relationship within the framework of a "Nostratic macrofamily". Emphasis is placed upon traditional methodologies such as the Comparative Method and Internal Reconstruction.

Walter de Gruyter, Inc., 200 Saw Mill River Road, Hawthorne, NY 10532, USA

MOTHER TONGUE

JOURNAL OF THE ASSOCIATION FOR
THE STUDY OF LANGUAGE IN PREHISTORY

AIM AND SCOPE

The Association for the Study of Language in Prehistory (ASLIP) is a nonprofit organization, incorporated under the laws of the Commonwealth of Massachusetts. Its purpose is to encourage and support the study of language in prehistory in all fields and by all means, including research on the early evolution of human language, supporting conferences, setting up a databank, and publishing a newsletter and a journal to report these activities.

MEMBERSHIP

Annual dues for ASLIP membership and subscription to *Mother Tongue* are US $25.00 in all countries except those with currency problems. For membership information, contact:

Anne W. Beaman, *Secretary*
Association for the Study of Language in Prehistory
P.O. Box 583
Brookline, MA 02146 U.S.A.

Design and Production:

SIGNUM
DESKTOP PUBLISHING
88-B QUEEN STREET
CHARLESTON, SC 29401-2427
(803) 720-8531